User Report

Solicitor General Canada
Ministry Secretariat

R. Karl Hanson
Brian Cox
Carolyn Woszcsyna

SEXUALITY, PERSONALITY AND ATTITUDE
Questionnaires for Sexual Offenders: A Review

1991-13

This working paper was prepared under contract for the Corrections Branch, Ministry of the Solicitor General of Canada and is made available as submitted to the Ministry. The views expressed are those of the authors and are not necessarily those of the Ministry of the Solicitor General of Canada.

This document is available in French. Ce document de travail est disponible en français.

Supply and Services Canada
Cat. No. JS4-1/1991-13
ISBN: 0-662-58615-8

93-00980/

Table of Contents

Introduction . . . 1
Selection of measures. . . . 7

Omnibus measures

Thorne Sex Inventory . . . 9
Multiphasic Sex Inventory . . . 13

Sexual history and interest measures

Clarke Sex History Questionnaire . . . 24
Sexual Interest Card Sort . . . 27
Sex Fantasy Questionnaire . . . 29
Sexual Experiences Survey . . . 30
Aggressive Sexual Behavior Inventory . . . 33
Self-Reported Propensity to Rape . . . 35

Child molesting attitude measures

Abel Becker Cognitions Scale . . . 36

Rape attitude measures

Burt's rape attitude scales. . . . 38
Attitudes Towards Rape Scale . . . 46
Rape Empathy Scale . . . 49
Rape Supportive Attitude Scale . . . 51
Rape Attitude Scale . . . 52

Sex role measures

Attitudes Towards Women . . . 54
Bias in Attitude Survey . . . 58
Index of Sex-Role Orientation . . . 59
Sexism Scale . . . 60
Attitudinal Inventory . . . 61
Hypermasculinity . . . 62
Heterosexual Relationships Scale . . . 64

Sexual dysfunction

Derogatis Sexual Functioning Inventory . . . 65
Negative Attitudes Towards Masturbation . . . 67

Other measures

Hostility Towards Women . . . 69
Personal Construct Assessment . . . 71
MMPI . . . 72

Conclusions and future directions . . . 78

References . . . 83

Introduction

This report reviews sex-related self-report measures that are potentially useful in the assessment of sexual offenders. Measures of sexual attitudes and sexual functioning are typically part of the comprehensive assessments of sexual offenders. Other factors important in such assessment are criminality, violence, substance abuse, and mental health problems. Phallometric assessments of sexual offenders have long been identified as contributing important information; however, these measures have already been reviewed elsewhere (e.g., Quinsey & Earls, 1990). Instead, this report focuses only on self-report measures of sexuality. Questionnaire measures, should they prove to be reliable and valid, have the distinct advantage of being inexpensive and convenient to administer. Although not a sexuality measure _per se_, the review also covers research on the MMPI because it has been so widely used in the assessment of sexual offenders.

Although numerous sex scales have been proposed, there have been few evaluative reviews of these measures. Hall (1990), Murphy (1990), and Laws (1984) have reviewed some sex history and sex attitude measures that could be used with sexual offenders; sexuality questionnaires were not, however, a specific focus of these reviews. Salter (1988) presented several sex measures for sexual offenders, but provided little evaluative review; and Derogatis (1980) reviewed sex role and sexual dysfunction measures, but his focus was on sexual dysfunction, not sexual deviance. General collections of sexuality scales have been presented by Davis, Yarber, and Davis (1988) and Schiavi, Derogatis, Kuriansky, O'Connor and Sharpe (1979).

The merits of particular measures depend, in part, on the applications for which they are being used. The perspective adopted in this report is that of risk management. This perspective assumes that the goal of assessment is to identify factors related to risk for sexual offending, and the goal of treatment is to change those risk factors amenable to change. Sexual offenders may have many problems worthy of treatment in their own right, such as depression or compulsive gambling; however, the focus of the present review is on the value of the measures as risk indicators for sexual recidivism.

The measures reviewed were selected because there was either research or theory linking the constructs assessed to sexual offending. Most theories of sexual offending emphasize multiple constructs or factors (Canada, 1990; Finkelhor, 1984; Quinsey, 1984, 1986). Some of these factors include deviant sexual interests, attitudes supportive of sexual offending, deficient sexual knowledge, sexual dysfunction, opportunities provided by exposure to potential victims, and the disinhibiting effects of deviant role models (both real and imagined). An initial consideration in the evaluation of sexual offender measures is the strength of the theory guiding the development of the scale. Clinicians and researchers would benefit from using measures that are consistent with current expert opinion (see Laws, 1989; Marshall, Laws & Barbaree, 1990). For example, global measures of self-esteem or adjustment within treatment programs appear to be unrelated to subsequent recidivism (Hall, 1988; Hanson, Steffy & Gauthier, 1990a). In general, the measures that appear to have the greatest utility are those that are closely related to the offending behaviour.

Once relevant constructs are identified, the next consideration is the reliability of the specific measures. Reliability refers to the "consistency of scores obtained by the same persons when reexamined with the same test on different occasions, or with different sets of equivalent items, or under other variable examining conditions" (Anastasi, 1982, p. 102). Reliability can be assessed in several different ways. One common method of assessing reliability is through repeated administrations of the same test, which is referred to as stability or test-retest reliability.

Internal consistency measures can also be used to assess reliability from single administrations of multi-item measures. Classical measurement theory assumes that scales assess some underlying latent construct. The response to each test item is influenced by this latent construct plus error. Combining items into scales improves reliability because additional items contribute additional "true score" variance, whereas the "error variance" is cancelled out (by definition). Consequently, the reliability of a test can be estimated by examining the correlations between test items. This can be accomplished through corrected split-half correlations (using the Spearman-Brown formula, see any standard psychometric text, e.g., Ghiselli, Campbell & Zedeck, 1981). Split-half correlations, however, are influenced by how the items are split. Consider a hypothetical test that contained the following four questions: 1) do you worry a lot? 2) do you often worry? 3) do you eat a lot of chocolate cake? 4) do you eat chocolate cake often? If questions 1 and 2 were combined and correlated with questions 3 and 4, the resulting correlation would be expected to be close to zero. If questions 1 and 3 were combined and correlated with questions 2 and 4, the correlation would be close to one. Consequently, rather than using split-half reliability measures, it is preferable to use formulas that estimate the average split-half correlations from all possible splits. These equations are KR-20 (Kuder & Richardson, 1937), for use with dichotomous items, and Cronbach's Alpha (Cronbach, 1951), for use with continuous or ordinal items (Ghiselli et al., 1981).

Factor analysis is another technique that has commonly been used to evaluate the internal consistency of measures. A few comments on factor analysis are in order since it is a technique widely misunderstood and misused. Factor analysis attempts to identify the underlying latent dimensions within a set of correlations (see Gorsuch, 1983, for a highly recommended introduction to factor analysis). Commonly, it is used to determine whether a scale contains only one or several dimensions. However, factor analysis does not provide one solution, but, instead, generates an infinite number of solutions that are all equally correct mathematically. Selecting the most plausible factor analytic solution requires making some assumptions about the expected solution. The two most difficult of these decisions are 1) determining the number of factors to extract and 2) deciding how these factors should be rotated. Computer programs often make these decisions using default options, giving the naive users the false impression that the default solution is the correct solution. In many cases, however, the default solution is not correct.

In many statistical packages (e.g., SPSSX, BMDP), the default option for determining the number of factors to extract is based on the number of factors whose unrotated eigenvalues are greater than 1.0 - the K1 rule (Kaiser, 1960). This method, however, is known to overestimate the number of factors to extract and to create spurious multidimensionality in scales (Zwick & Velicer, 1986). Zwick and Velicer (1986) report that the K1 rule (i.e., eigenvalue greater than 1) is sufficiently inaccurate that it "should not be used" (p.441). The most accurate techniques for determining the number of factors to extract require computer programs that are not in common use. However, Zwick and Velicer (1986) found the Scree Test (Cattell, 1966) to be reasonably accurate and easily conducted. The Scree Test involves examining the plot of the eigenvalues of the unrotated factors and retaining only the number of factors for which there are eigenvalues that are noticeably greater than the next largest eigenvalue. It is called a "scree" test because it requires a judgement as to where the "cliff" separates the "mountain" from the rubble in the "scree slope" below. (A "scree slope" is a geological term for the material that collects at the bottom of cliffs.) Although the Scree Test is more accurate than the K1 rule, most of the researchers who factor analyzed their sex scales seemed to uncritically adopt the K1 rule. The consequence of this error is that many unidimensional scales will appear multifactorial.

Once factors are extracted, the common default rotational procedure is Varimax. This procedure assumes that the underlying dimensions are not correlated with each other (orthogonal). Although this is a reasonable assumption in some circumstances, it does not appear appropriate when factor analyzing scale items that were intended to represent a single underlying construct (e.g., pro-feminist attitudes). The alternative is to allow the resulting factors to correlate with each other by opting for an oblique rotation procedure. Almost all the factor analyses of the sex scales used Varimax rotations; however, in many cases the resulting factors were significantly correlated with each other, which questions the assumption that the identified factors were really orthogonal.

None of the factor analyses reported on the sex scales attempted to assess the adequacy of the final solution using confirmatory factor analyses such as LISREL (Joreskog & Sorbom, 1986). Although there is debate about the best criteria to use (March, Balla & McDonald, 1988; Muliak et al., 1989), these indices can, nonetheless, provide some guidance when choosing between alternate factor analytic solutions.

Both Cronbach's Alpha and factor analysis are concerned with internal consistency, which is one type of reliability. Internal consistency is a valuable characteristic of scales, but it is not essential for many types of measures. In particular, well constructed screening instruments, such as symptom checklists and history measures, need not have high intercorrelations between the items. When the responses to individual items are meaningful and a "total score" is not calculated, then it does not matter whether the items correlate with each other. If an offender, for example, indicated on a questionnaire that he had rubbed his penis against two girls aged 12 or younger, the offender's response is important, regardless of his responses to other items. The examiner would want to know whether the offender understood the question and provided the response he had intended, but the examiner would not be particularly concerned about whether the question correlated with other items. Similarly, a sexual dysfunction measure may contain items about premature ejaculation and impotence. Both of these items are appropriate on a measure of sexual dysfunction, although the correlation between these items would be expected to be negative. The reliability of such "checklist" screening measures is more appropriately assessed by test-retest measures than by internal consistency measures.

There are several different approaches to evaluating the validity of the sex measures. Since the major concern of this review is risk management, the strongest test of a measure is its ability to predict future sexual offending. The research to date, however, has been insufficient to establish the predictive validity of any of the measures reviewed. There have been few recidivism studies, and the results have tended to be inconclusive (Hall, 1990). Offense history is a predictor of recidivism (e.g., Quinsey, 1977, 1984, 1986; Hall, 1988; Hanson, et al., 1990a); consequently, it is quite likely that questionnaire sexual history measures could predict recidivism. Such studies have yet to be completed. It also remains questionable whether the typical self-reported offense histories would have any predictive power over offense histories as recorded in official records. Nevertheless, under conditions of unusual confidentially, offenders' can report information with considerable clinical utility. Abel et al. (1987), for example, found that some sexual offenders reported hundreds of diverse offenses when assured that such disclosures could not be used against them.

Another strategy for establishing validity is to compare known groups of sexual offenders with nonoffenders. When differences are found, then the characteristic assessed receives some support as being related to sexual offending. The continual problem with such correlational research, however, is that it is frequently possible to imagine how the observed correlations could result from spurious factors, such as social class, reading ability, or the effects of incarceration. Consequently, the adequacy of the comparison groups needs to be closely evaluated, and multiple comparison groups are required. The comparison groups should contain subjects from populations for whom the test is intended. For example, if it is intended to be used with criminals, then it should at least be validated with criminals. One useful design is to compare the index groups (e.g., pedophiles) to another deviant groups (e.g., property offenders or non-pedophilic sexual offenders) as well as to a non-deviant groups (e.g., nonoffenders matched for age and education). Although the results of such studies are never definitive, the cumulative results of such comparisons can be highly informative.

Even when a measure does not distinguish between sexual offenders and nonoffender comparison groups, some authors have argued that the measure might assess a factor associated with sexual offending (e.g., Alford & Brown, 1985; Burt, 1983). Such arguments suggest that the attribute in question, in this case sexist attitudes, functions differently for sexual offenders than for nonoffenders (Alford & Brown, 1985; Burt, 1983). These arguments seem strained, but are not entirely indefensible. It is possible to argue, for example, that an attribute interacts (in the statistical sense) with other variables to contribute to sexual offending. Sexual offenders and nonsexual violent offenders may have an equal amount of alcohol problems, but it is still plausible to argue that alcohol abuse contributes to sexual offending - but only in those individuals otherwise predisposed to sexual violence. In support of this general view, Malamuth (1986) has shown that interactions between various factors are stronger predictors of self-reported sexual aggression than the direct effects of these same factors. However, if researchers observe no group differences and still want to argue that the measure is related to sexual offending, then it is incumbent upon the researchers to specify and test the hypothesized interaction. If such further analyses are not conducted, then the most parsimonious (and plausible) conclusion is that the attribute is not related to sexual offending or that the measure is inadequate. It is also important to remember that sexual offenders and nonoffenders may have attributes that are offensive or problematic (e.g., racist or sexist views, depression) that are worthy of intervention even though these attributes may have little connection to sexual offending.

The most common method for assessing the validity of a measure is by correlating it with other conceptually related and unrelated measures. The convergent and discriminant validity of the measure can then be analyzed in the resulting multitrait-multimethod matrix (Campbell & Fiske, 1959; Schmitt & Stults, 1986). A measure of sexual aggression, for example, should correlate with other sexual aggression measures, but it should not correlate with measures of anxiety or social desirability. General comparisons with other measures were frequently used in determining the validity of the sex measures included in this review. Complete multitrait-multimethod matrices, however, were rarely analyzed or reported.

A persistent concern when assessing sexual offenders is the problem of denial. Offenders face strong social pressures to deny or minimize their deviant sexuality (Taylor, 1972). Clinical assessments are not anonymous and offenders who appear at risk for reoffending can face serious consequences related to sentencing, conditional release, and child custody. Offenders who deny their sexual offenses need not appear to be minimizing problems on standard measures of social desirability, such as the MMPI Lie scale or the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960). It is easy for offenders to admit to the minor faults associated with social desirability scales while failing to acknowledge their criminal sexual acts. The only adequate method of assessing denial is by comparing the acts that offenders self-disclose to the acts recorded by more credible sources (e.g., victim reports, police records). Since offenders who do not admit to their offenses are poor experimental subjects, they are often excluded from studies. Consequently, much of what we know about sexual offenders is based on select samples. Most of the validation studies included in this review examined either admitting sexual offenders or nonoffender samples for whom there were no consequences to their responses. It is not clear that measures developed on such samples have any merit in applied assessment contexts.

One approach to avoiding offenders self-incriminating themselves is to use techniques that guarantee anonymity, such as the randomized response technique (Chaudhuri & Mukerjee, 1988; Fox & Tracy, 1986). This technique involves camouflaging subjects' answers by introducing systematic error. One of the several methods involves instructing the subject to respond "yes" 10% of the time, "no" 10% of the time, and to answer truthfully 80% of the time. The yes, no, and truthful responses are determined by some random method such as dice or party spinner. Finkelhor and Lewis (1988) used a randomized response technique in a telephone interview investigating the prevalence of child molesters in Los Angeles. They found that the rates were strongly influenced by which of two "safe" questions were used as foils. The difference between the alternate administrations was sufficiently large (4%

versus 17%) to questions the validity of the technique used in this study.

Another randomized response technique is currently being used at the model treatment program at Atascadero State Hospital in Sacramento, California (David Day, personal communication). The initial results from the Atascadero measures indicated that the usefulness of these measures is hindered by their low statistical power. (The power of the test varies according to the amount of error introduced into the measurement). Even with samples sizes of 150, differences between 10% and 20% are usually not statistically significant, which suggests limited utility program evaluation measures. Randomized response techniques obviously cannot be used for individual assessments, but are potentially useful in large scale surveys involving hundreds of subjects.

Another strategy for circumventing self-presentation biases is to use skill or competency measures. On skill measures offenders are not asked to self-disclose; rather, they are asked to provide their best responses. Such tests aim to identify deficits rather than simply differences. Skill deficit measures for sexual offenders are just beginning to be systematically studied, but the initial results are promising (Lipton, McDonel & McFall, 1987). One valuable strategy is to develop program specific skill measures, as has been used in relapse prevention programs (e.g., Miner, Day & Nafpaktitis, 1989). This approach involves specifying the knowledge and skills that the offenders are expected to learn in the treatment program, and then constructing tests to assess achievement of these treatment goals. These measures can be as simple as multiple choice tests about the treatment model; more complicated measures involve challenges to provide coping responses to high risk situations (Miner et al., 1989).

The criteria for evaluating skill measures are essentially the same as the criteria used for evaluating other measures. For program specific measures, there is the added concern of determining acceptable minimal levels, i.e., passing scores. Until recidivism data is collected, the only reference for determining minimum standards is the theory guiding the treatment program. The typical comparison groups cannot be used since programs often teach many new skills that would not be expected to be known by untreated subjects (e.g., relaxation techniques, covert sensitization).

Selection of measures

Several criteria guided the selection of sex measures for this review. The most important factor was the measure's possible value as a risk indicator. Considering the lack of comprehensive evaluative research, the selection of measures was primarily guided by the opinions of prominent clinicians regarding the utility of various tests.[1] The inclusion of a measure in this review should not be considered an endorsement; a test could be included in the review simply because it had been used with sexual offenders, or there had been a suggestion that it may be associated with sexual offending. Although the review was comprehensive, omissions are inevitable. We would be most interested in hearing from readers who know of relevant scales that have been overlooked. The following reviews of the specific scales need not be read in any particular order. Occasionally, however, readers are referred to other sections of the review for additional comments about related measures.

---

[1] We would like to thank the following for providing valuable assistance for this review: Gene Abel (Atlanta, GA), David Day (Sacramento, CA), Bruce Etches (Maple Ridge, B.C.), David Finkelhor (Durham, NH), Art Gordon (Saskatoon), Grant Harris (Penetanguishene), Sharon Hodkinson (Kingston), Steve Kahn (Halifax), Richard Laws (Edmonton), Janice Marques (Sacramento, CA), Gjylena Nixhipi (Brampton), Anne Pawlak (Ottawa), Isabelle Pronovost (Quebec), Vernon Quinsey (Kingston), Carson Smiley (Abbotsford, B.C.), Lana Stermac (Toronto), Lee Studer (Edmonton), and Sharon Williams (Kingston).

Thorne Sex Inventory (SI)

Intended use

The SI was designed as a clinical screening instrument for the assessment of identified and/or potential sexual offenders (Thorne, 1966a,b). It was intended to be used with an interview in which responses to critical items - "stop" items - were probed. The subscales address sexual attitudes, sexual dysfunction, and sexual deviations.

Description

The SI is composed of 200 true/false questionnaire items. Forty of the questions are considered "subtle-indirect" items; for example, one subtle measure of homosexuality is "It makes me uncomfortable to be the centre of attention." The remaining 160 items are "obvious-direct" items, e.g., "I have been arrested for a sex offense." Thorne (1966a) reported that the items were developed on the basis of "extensive empirical experimentation with interview methods and questionnaire items" (p. 367), using a diverse sample of sexual offenders. The items are arranged into nine subscales based on theory and factor analyses (Thorne, 1966a, 1966b). Each of the subscales will be reviewed separately.

A. Sex Drive and Interest

The 21 items (3 true, 18 false) of this scale address liberal versus conservative sexual values. The items, however, are quite diverse and are not limited to sexuality. Example items include "Virginity is a girl's most valuable possession," "I often pray," and "Body odors are disgusting." Thorne (1966a) reported that these items loaded on a common factor, and Allen and Haupt (1966) reported a three month stability coefficient of .83 (intraclass correlation). A mixed group of sexual offenders scored more liberal on this measure than groups of nonsexual offenders, but scored more conservatively than college males (Haupt & Allen, 1966). Howells and Wright (1978) found that aggressive sexual offenders were similar to mentally disordered aggressive offenders on this scale. Cowden and Pacht (1969) similarly found that sexual and nonsexual offenders were not significantly different on this scale.

B. Sexual Maladjustment and Frustration

The 21 items of this scale (17 true, 4 false) address self-reported unhappiness with heterosexual relationships. Example items include "Thoughts about sex disturb me more than they should," and "I have never had many dates." Thorne (1966a) reported that these items load on a common factor, and Allen and Haupt (1966) reported a three month stability coefficient of .79 (intraclass correlation). A mixed group of sexual offenders reported more sexual maladjustment than did nonsexual offenders or college males (Haupt & Allen, 1966). Howells and Wright (1978) similarly found that aggressive sexual offenders scored higher than other mentally disordered aggressive offenders on this scale. Cowden and Pacht (1969) also found that sexually deviated sexual offenders scored more highly on this scale than did either nonsexual offenders or criminals convicted of a sexual crime but not judged to be sexually deviated.

C. Neurotic Conflict Associated With Sex

These 21 items (all true) address subjective distress and negative personal history. An example item, "I often feel sad and blue." The scale would be more appropriately entitled Neurotic Conflicts since only one item makes specific reference to sexual problems. Since the scale does not assess sex related attitudes or problems, it will not be considered further.

D. Sexual Fixations and Cathexes

The 37 items (4 true, 33 false) of this scale address a variety of sexual preferences and sexual deviations (e.g., paedophilia, voyeurism, anal eroticism). The items are not intended to be an internally consistent scale, but are a screening checklist for specific sexual deviations. Allen and Haupt (1966) reported a three month stability coefficient of .75 (intraclass correlation). A mixed group of sexual offenders reported the same overall score on this measure as did nonsexual offenders and college males (Haupt & Allen, 1966). Similarly, Howells and Wright (1978) found that it did not distinguish between sexual and nonsexual offenders.

E. Repression of Sexuality

The 19 items (12 false;7 true) of this scale address openness to sexual experience. Example items include "I get sexually excited very easily" and "My religious beliefs are against sex." Thorne (1966a) reported that these items load on a common factor, and Allen and Haupt (1966) report a three month stability coefficient of .79 (intraclass correlation). A mixed group of sexual offenders reported more repression than college males but somewhat less repression than nonsexual offenders (Haupt & Allen, 1966). Other researchers, however, found that it did not distinguish between sexual and nonsexual offenders (Cowden & Pacht, 1969; Howells & Wright, 1978).

F. Loss of Sex Controls

These 18 items (17 true, 1 false) address deviant sexual impulses. Example items include "Afraid of what I might do sexually," and "I have been accused of molesting a child." Thorne

(1966a) reported that some of these items load on a common factor, but that other items were added because they were conceptually related. Allen and Haupt (1966) report a three month stability coefficient of .76 (intraclass correlation). A mixed group of sexual offenders scored well above nonsexual offenders (Cowden & Pacht, 1969; Haupt & Allen, 1966; Howells & Wright, 1978) and college males (Haupt & Allen, 1966). Such a finding is not particularly surprising since many of the items are concerned with being in trouble because of sexual problems.

G. Homosexuality

These 20 items (18 true, 2 false) are mainly indirect measures of potential homosexual interest. An example item: "I feel nervous with the opposite sex." Thorne (1966a) reports that these items load on two conceptually related factors. Allen and Haupt (1966) report a three month stability coefficient of .80 (intraclass correlation). A mixed group of sexual offenders scored well above (one standard deviation above) nonsexual offenders and college males on this scale (Haupt & Allen, 1966). Cowden and Pacht (1969) similarly found that it distinguished sexual and nonsexual offenders. In Haupt and Allen's (1966) study, however, the difference between homosexual and heterosexual sexual offenders on this scale was less than the difference between sexual and nonsexual offenders, which raises questions about the construct validity of this measure. Howells and Wright (1978) found that it did not distinguish between sexual and nonsexual offenders.

H. Sex Role Confidence (SRC)

These 12 items (4 true, 8 false) address self-reported unhappiness with heterosexual adjustment. It correlated .50 with the Sexual Maladjustment and Frustration scale (Haupt & Allen, 1966) and contains some of the same items. Thorne (1966a) reported that SRC items load on one factor. Allen and Haupt (1966) reported a three month stability coefficient of .79 (intraclass correlation). A mixed group of sexual offenders reported less sex role confidence on this scale than nonsexual offenders and college males (Haupt & Allen, 1966). Howells and Wright (1978) and Cowden and Pacht (1969), however, found that it did not distinguish between sexual and nonsexual offenders.

I. Promiscuity

These 18 items (17 true, 1 false) address low "superego" functioning (i.e., moral functioning). Most, but not all, of the items are sex related. Example items include "I have paid money for sex," and "I believe in taking my pleasures where I find them." Thorne (1966a) reported that these items loaded on two conceptually related factors. Allen and Haupt (1966) report a three month stability coefficient of .75 (intraclass correlation). There were no differences on this measure between a mixed group of sexual offenders, nonsexual offenders and college males (Haupt & Allen, 1966; see also Cowden & Pacht, 1969; Howells & Wright, 1978).

Overall summary and evaluation

Apart from the research of Thorne and his associates (Allen & Haupt, 1966; Haupt & Allen, 1966; Thorne & Haupt, 1966), the SI has received little research attention and does not appear to be widely used in current clinical settings. The lack of interest can partially be attributed to changing perspectives on sexual offenders. The theoretical model guiding the SI emphasized the importance of sexual dysfunction and ego-dystonic sexual impulses. Homosexuals were included as sexual offenders and the SI scales were not intended to distinguish between different types of sexual offenders, such as rapists and paedophiles. An additional problem with many of the scales is the heterogeneity of the scales. The inclusion of "subtle" items makes the interpretation of many of the scales ambiguous.

The SI's strength is as a loose collection of questions that can function as a screening instrument to accompany further interview assessment. In its current form, however, many of the questions are superfluous. Consequently, it is not surprising that subsequent researchers have used the Thorne SI primarily as a source of questions to be incorporated into new scales (e.g., Briere & Runtz, 1989; Eysenck, 1971; Whalley & McGuire, 1978).

Available from:

There does not appear to be a manual for the SI. All the items (except the Sexual Fixations scale) are available in Thorne (1966a). Most of the available psychometric information on the scale is in the same issue of the Journal of Clinical Psychology, 22, 1966, pp. 367-407.

Multiphasic Sex Inventory (MSI)

Intended use

The MSI (Nichols & Molinder, 1984) was constructed to assess a range of sexual characteristics of identified sexual offenders, including child molesters, rapists, exhibitionists, and other paraphiliacs. The subscales assess attributes related to sexual attitudes, sexual history, sexual knowledge, sexual dysfunction, and denial/defensiveness concerning sexual behaviour. It is not appropriate for nonsexual offenders.

Description

The MSI is a 300 item true/false questionnaire. Offenders would be expected to take one to 1.5 hours to complete it. Earlier versions were constructed as far back as 1977, although the current version was completed in 1984. The form contains both validity scales and content scales. The subscales are reviewed separately.

A. Parallel Items (a subscale)

The MSI contains two sets of parallel items: 15 items duplicated from the MMPI, and 19 from the Thorne Sex Inventory. The aim of the above scales is to screen for careless or random responding. This method of checking for random responding has strong face validity. The utility of these scales, of course, require the concurrent administration of either the MMPI or the Thorne Sex Inventory.

B. Social Sexual Desirability

This scale contains 35 items (15 true, 20 false) that aim to assess the offenders' denial and/or minimization of normal sexual interests. Typical items include "I am not interested in sex matters like most men seem to be," and "There are times that I laugh at a dirty joke" (scored if answered "false"). Research on an earlier 14 item version of the scale (reported by Nichols and Molinder, 1984) suggested it had adequate internal consistency (KR-20 = .71) and moderate three week stability ($\underline{r}$ = .64). They also found that a group of offenders who completed treatment reported more normal sexual interests than a separate group of offenders who had not attended treatment. There is little information in the MSI manual, however, concerning the revised 35 item scale, other than basic descriptive statistics (a mean and standard deviation) from a group of 40 men attending a night school college course.

Simkins et al. (1989) used the MSI in a study of child molesters attending a family oriented treatment program. Most of the men were intrafamilial offenders. Simkins et al. found that the three month stability of the 35 item Social Sexual Desirability scale was .84. They also found that scores did not change between initial assessment and nine months into treatment. They did find, however, that offenders who were willing to report normal sexual interests on this scale were also the offenders who showed the least denial of their offenses, and who made the most progress in treatment (according to therapists ratings).

Gillis (1991) reported high internal consistency for the 35 item scale (KR-20 of .87). The scale, however, did not distinguish between the rapists, property offenders, and violent nonsexual offenders assessed in his study.

These preliminary results suggest that the scale has adequate internal consistency and stability. The construct it intends to assess appears conceptually relevant for research and assessment of child molesters. As well, there is some research supporting its utility with this population. The utility of this scale with other paraphilic groups remains unknown.

C. Sexual Obsessions

This scale contains 20 items (all true) intended to assess high levels of sexual fantasy and activity, e.g., "It seems that everything I do and everywhere I go I am constantly thinking about sex." Nichols and Molinder (1984) interpreted moderate scores as evidence of excessive and/or intrusive sexual drive. They interpret very high scores as malingering. The manual contains no psychometric information on the scale, other than the mean and standard deviation for 40 college students. They do, however, report moderate internal consistency (KR-20 of .65) and stability ($\underline{r}$ = .70) for a preliminary 12 item version of the scale. They also reported that their treatment group reported more sexual obsessions than the untreated offenders on the 12 item scale.

Simkins et al. (1989) found the three month stability of the 20 item scale to be high ($\underline{r}$ = .80) for a group of offenders in treatment. Simkins et al. also reported that the scale did not change during the course of treatment and was unrelated to treatment outcome. They did find, however, that the offenders who selected both female and male victims and/or selected both intra- and extra-familial victims reported higher sexual obsessions than the child molesters who selected only one type of victim.

The scale has strong face validity and is likely to be reasonably internally consistent and stable. Its utility has yet to be established.

D. Child Molester Lie Scale

This scale contains 13 items (3 true, 10 false) intended to assess the extent to which identified child molesters deny/minimize their sexual activities with, and sexual interest in, children. The items are direct admissions of characteristics that most child molesters are likely to have, e.g., "Sometimes I am sexually attracted to children" (scored if answered "false"). This scale is simply the reverse scoring of 13 items from the 39 item Child Molest content scale to be discussed later.

No reliability data was presented for this scale. Simkins et al. (1989) reported that it did not change as a result of treatment; however, they did find that it was correlated with clinician ratings of positive treatment outcome and denial. Nichols and Molinder (1984) similarly reported that child molesters who completed treatment scored lower on this scale (admitted more deviance) than did an unselected group of untreated child molesters. Simkins et al. (1989) found that offenders who admitted to victims both inside and outside the family, and to both female and male victims, scored lower on this scale than did the other child molesters in their study.

E. Rape Lie Scale

This scale contains 13 items (1 true, 12 false) that aim to assess denial/minimization in identified rapists. As with the Child Molester Lie Scale, the Rape Lie Scale is simply the reverse of items in the 28 Rape Scale that are likely to be true for most rapists, e.g., "I have attempted rape or raped at least once" (scored is answered "false").

Gillis (1991) reported high internal consistency for this scale (KR-20 of .87) within a mixed group of rapists and nonsexual criminal controls. As with the Child Molest Lie Scale, Nichols and Molinder (1984) reported that scores were lower for rapists who completed treatment than for an unselected group of untreated rapists.

F. Incest Lie Scale

This scale is simply the reverse of four items from the Incest Scale. No reliability or validity data were available.

G. Exhibitionism Lie Scale

This scale is the reverse of six items from the Exhibitionism Scale. No reliability or validity data were available.

Comment on Lie scales: The lie scales are intended to measure the offenders' willingness to admit to sexual offenses and deviations. The utility of all the lie scales requires accurate information concerning the offenders' actual offenses. Without such information, the results are uninterpretable. It cannot distinguish between offenders who are denying specific sexual activities and those offenders (and nonoffenders) who have never committed the acts.

H. Cognitive Distortion and Immaturity (CDI)

The 21 items (19 true, 2 false) of the CDI are intended to assess the extent to which the offender adopts a "victim stance" in relationship to his present offense, e.g., "In some ways I was used by the person who reported me." The items, however, appear quite heterogeneous: there are items that address identification with children, feeling generally misunderstood, delayed development of sexual interest, and feeling unloved by one's parents. Considering the paucity of reliability and validity data, it is difficult to determine what the scale is assessing. Nichols and Molinder explicitly fail to report internal consistency estimates on the erroneous assumption that unequal item difficulties preclude the usefulness of formuli such as KR-20. Nichols and Molinder are partially correct in that item difficulty can decrease inter-item correlations for dichotomous items, thus deflating internal consistency estimates. For most practical purposes, however, the decrease is not significant. The usual correction involves the use of tetrachoric coefficients of correlation for dichotomous items based on theoretically continuous underlying constructs (see any standard psychometric text, e.g., Ghiselli et al., 1981). The only data presented in the test manual are the mean and standard deviation based on a small sample (31) of untreated sexual offenders. Gillis (1991), however, reported an internal consistency coefficient of .73 (KR-20); Simkins et al. (1989) reported a three month stability coefficient of .71.

Simkins et al. (1989) found that child molesters who offended both inside and outside the family scored higher than other offenders. Low scores predicted positive outcome in their family oriented treatment of child molesters, but the scores did not change with treatment. In contrast, Miner, Marques, Day and Nelson (1990), found that the scores from this scale decreased during their cognitively oriented relapse prevention treatment program. The posttreatment scores of child molesters were also lower than the scores for a matched, randomly assigned control group.

Gillis (1991) found that the CDI failed to distinguish between rapists, nonsexual nonviolent criminal, and nonsexual violent criminals. He did find, however, that the scale showed moderately strong correlations with other measures typically used to assess sexual offenders, such as the Hostility Towards Women Scale (Check, 1985), The Buss-Durkee Hostility Inventory (Buss & Durkee, 1957), and some of the subscales of Burt's (1980) Rape Myths Scale. Kalichman and his colleagues found the CDI distinguished between MMPI derived subgroups of rapists (Kalichman, Craig et al., 1989; Kalichman, Szymanowski, McKee, Taylor & Craig, 1989).

The heterogeneity of the items and the lack of adequate psychometric analysis makes the interpretation of this scale problematic. Some of the items may be related to sexual offending (to child molesting more than to rape), but basic psychometric research (e.g., item-total correlations, factor analyses) need to be conducted before the scale can be considered to be clinically useful.

I. Justifications Scale

The 24 items of this scale (all true) address various justifications that sexual offenders may use to explain their offenses, e.g., "My sexual offense occurred because of the stresses in my life." Other justifications refer to alcohol, victims, and other external circumstances. No reliability data were presented in the manual, and it is questionable whether the scale should be considered to be unidimensional. In other words, it is not clear that sexual offenders who use one justification are more likely to use other justifications as well. The only data in the manual are the mean and standard deviation from a group of 31 untreated sexual offenders.

Simkins et al. (1989) found the three month stability of the scale to be .78. It was weakly correlated with positive treatment outcome, and did not change with treatment. Offenders in their study who showed less denial tended to report more justifications.

Miner et al. (1990) found that the Justifications scale decreased with treatment. As well, the posttreatment scores of child molesters were lower than the scores for an untreated control group. Similar differences were not noted for rapists, although there was an insufficient number to provide an adequately powerful test.

J. Treatment Attitude Scale

The eight items (3 true, 5 false) of this scale assess the extent to which offenders consider themselves to have a sexual problem justifying treatment, e.g., "I need help because I am not able to control my sexual behaviour." Again, the only data available in the manual are the mean and standard deviation from 31 untreated sexual offenders. Simkins et al. (1989) found the three month stability of the scale to be .79. It was correlated with positive treatment outcome. As well, scores were highest for offenders who showed little denial and who offended both inside and outside the family.

The scale has strong face validity, although it is unclear whether it is better than simply directly asking the offender whether he feels that he has a sexual problem that requires treatment.

K. Child Molest Scale

The 39 items (24 true, 11 false) of this scale are direct questions concerning sexual interest and sexual activity with children, e.g., "I have been attracted to boys sexually." As well, there are an additional four items that comprise an Incest subscale. The 39 item Child Molest scale is further subdivided into subscales entitled Fantasy (10 items), Cruising/grooming (10 items), Sexual Assault (9 items), and Aggravated Assault (6 items). The Sexual Assault scale includes items related to the actual commission of sexual offenses, and the Aggravated Assault contains items related to serious and/or repetitive sexual offending. Nichols and Molinder consider these subscales to represent the progression that leads to serious child molesting. The manual presented no psychometric information on the 39 item scale, although some information on an earlier version was presented. Nichols and Molinder found that child molesters who had completed treatment admitted to more offenses than did an unselected group of untreated child molesters. As well, they found the three week stability of the earlier version to be .91.

Simkins et al. (1989) found the three month stability of the revised 39 item scale to be .78. They also found that the scale was inversely related to clinician ratings of denial and positive outcome in treatment. As well, child molesters who offended both inside and outside of the family scored higher on this scale.

The four stage model implied by this scale is interesting, but it is unclear whether the items selected for this scale would be generally applicable to child molesters. For example, many incest offenders spend little time "cruising parks" looking for potential victims. Clinicians working with the scale should best consider it a checklist of individual items rather than an homogenous scale. No internal consistency data on this scale was available.

L. Rape Scale

The 28 items (21 true, 7 false) of the Rape Scale address rape interest and rape behaviour. The questions are direct, and, like the Child Molest scale, are further subdivided into subscales entitled Fantasy (8 items), Cruising (8 items), Sexual Assault (7 items), and Aggravated Assault (10 items). The subscales are considered to be stages in a progression leading up to serious sexual assaults. Again, the manual presented no information on the revised 28 item scale. It does report, however, that rapists who had attended treatment scored higher on an earlier 23 item version of the scale than did rapists who had not attended treatment.

Presumably the scale is intended to assess admission to already identified offenses.

Gillis (1991) found the internal consistency of the 28 item scale to be high (KR-20 = .88). He also found that rapists scored higher on the scale than comparison groups of violent and nonviolent criminal controls. The scale correlated moderately with the total score of Mosher's (1988a) Aggressive Sexual Behavior Inventory (*r* = .53).

M. Sado-Masochism

The ten items (all true) of this scale address sadistic sexual interest and behaviour, e.g., "I have fantasized about killing someone during sex." The title "Sexual Sadism" would be better for this scale since none of the items directly address masochistic sexual tendencies. Nichols and Molinder sometimes consider the scale part of the Rape scale, although they also include it in the Paraphilia scale, and discuss it separately (without presenting any psychometric data). Gillis (1991) reported moderate internal consistency for this scale (KR-20 of .72) and found that it correlated .38 with the MSI rape scale. Gillis did not find, however, that rapists scored higher on this scale than comparisons groups of violent and nonviolent criminal controls.

As with the Rape scale, the Sado-Masochism scale should best be considered a list of critical items that could alert clinicians to potential problem areas.

N. Exhibitionism

The 19 items (13 true, 6 false) of this scale address exhibitionistic interest and activity. As with the previous scales, it is subdivided into Fantasy (3 items), Cruising (4 items), Sexual Assault (9 items), and Advanced Assault (3 items). There is no psychometric information available in the manual or in any of other studies located.

O. Paraphilias

This scale is a collection of subscales related to Fetishism (9 items), Voyeurism (9 items), Obscene Phone Calls (4 items), Bondage and Discipline (6 items) and the 10 items from the Sado-Masochism scale previously discussed. The following is a sample item: "I have reached orgasm while secretly watching someone." The heterogeneity of the scale indicates that measures of internal consistency are probably inappropriate for this collection of questions. The manual reports high three week stability (*r* = .83). Simkins et al. (1989) found the three month stability to be predictably lower (*r* = .71). Simkins et al. found that the offenders who selected both male and female victims, and those who offended both inside and outside the family, scored higher on the Paraphilia scale than did those offenders with a more limited range of victims.

P. Sexual Dysfunction

This is a collection of items related to Sexual Inadequacies (8 items), Premature Ejaculation (4 items), Physical Disabilities (8 items), and Impotence (12 items). Example items include "I feel so foolish about climaxing so fast that I avoid women," and "Most of the time I cannot get an erection when I would like to have sex." Rather than an internally consistent scale, these items are intended as a screening checklist, with endorsement of critical items indicating the need for further assessment. The manual presents no psychometric information on this version of the scale. The three week stability of an earlier version was .80. The earlier scale, however, did not distinguish between sexual offenders and college students. Simkins et al. (1989) found the three month stability of the revised scale to be .77. Simkins found that this measure was unrelated to any of the other measures used in their study.

Q. Sex Knowledge and Belief

The 24 items (11 true, 13 false) in this scale assess knowledge of sexual anatomy and physiology, e.g., "A woman urinates through her clitoris." On an earlier version of the scale, Nichols and Molinder found that male college students scored about one point better than sexual offenders (19 versus 18). The internal consistency of the earlier version was low (KR-20 of .40), and the 21 day stability was marginal (r = .58). They present no psychometric data on the revised scale, but suggest that scores below 17 indicate the need for sex education. Simkins et al. (1989) found the three month stability of the revised scale to be .68. They also found that it was correlated with therapists' ratings of positive treatment outcome for child molesters, and that scores increased during treatment. The offenders who molested both boys and girls had higher scores than did the other child molesters in their study - a result that is difficult to interpret.

R. Sex History

These questions address other factors of offenders' backgrounds that could be important in the assessment of sexual offenders. These questions are not intended to be internally consistent scales, but rather function as screening checklists. The Sex Deviance Development items (11 items) cover sexual problems in childhood, including five questions on being a victim of childhood sexual abuse. The Marriage Development items (10 items) cover the adequacy of adult heterosexual adjustment, e.g., "I have never been in love." There are three items related to Gender Identity Development and six items dealing with Gender Orientation Development, which address, respectively, transsexualism and homosexual orientation. Finally, there is a list of 20 Sexual Assault Behavior items, which are duplicates from the Sexual Assault subscales from the Rape, Child Molest and Exhibitionism scale. No psychometric information is presented on these scales.

Other relevant research

In addition to the studies already mentioned (Gillis, 1991; Miner et al., 1990; Simkins et al., 1989), the MSI has also been used in Kalichman's research on MMPI clusters of rapists (Kalichman, Craig et al., 1989; Kalichman, Szymanowski et al., 1989). They found that the MSI Rape, Paraphilia, and Cognitive Distortion and Immaturity scales distinguished between five empirically derived groups of rapists. The interpretation of these differences, however, is hindered by many of the problems with the MSI already mentioned.

Simkins et al. (1989) subjected the MSI subscales to factor analysis, using the eigenvalue-greater-than-one rule for the number of factors to extract, and then subjecting the factors to a varimax rotation. The three factors that they extracted were not pure in that many items loaded on more than one factor; furthermore, it was difficult to identify common themes in the factors. This finding suggests that the individual scales are either multifactorial or that too many factors were extracted. Both alternatives are plausible. The lack of attention to internal consistency in the construction of the MSI scales makes it likely that the scales are heterogeneous. As well, the eigenvalue-greater-than-one rule is known to result in extraneous noise factors (Cliff, 1988; Zwick & Velicer, 1986). Simkins et al. do not provide sufficient information on the eigenvalues of the unrotated factors in order to examine whether the three factors were justified based on the Scree Test.

Simkins et al. (1989), however, found that the MSI was useful in identifying child molesters who were most likely to be rated as having positive therapy outcome. The MSI did not change during the course of their treatment, but it identified the offenders who expressed some motivation to address their admitted sexual problems. Simkins et al.'s (1989) overall evaluation of the MSI was that "[i]n many respects the MSI is a measure of denial" (p. 224).

Overall Summary and Evaluation

The MSI is an ambitious scale that attempts to assess a wide range of sexual characteristics of sexual offenders. Nichols and Molinder's basic assumptions about sexual offenders appear to be clinically relevant and plausible. Importantly, they recognize, and devote considerable discussion to, the issues of denial and minimization. The resulting scales, however, have serious limitations that compromise their utility.

Many of the problems with the MSI can be linked to ambiguity in the authors' measurement theory. Items were initially selected for their face validity in assessing the constructs of interest. Many of the items were sufficiently clear that affirmative responses to single items were highly interpretable, e.g., "I have molested a boy." As such, most of the MSI items can function as a checklist of critical items, any of which deserve attention if endorsed. The greatest strength of the MSI is that it provides a comprehensive list of important content areas that are relevant for sexual offenders.

The problem with the MSI begins when the items are combined into scales. Rather than combining items based on clearly defined latent constructs, most of the scales were constructed by simply grouping items into general categories. Consider the Paraphilias scale, which contains items related to fetishism, voyeurism, obscene phone calls, discipline, bondage, and sadism. All the items are relevant when assessing sexual offending. It is not clear, however, that the scale was intended to assess some unidimensional latent construct, such as "paraphilic tendencies." Similarly, on the Child Molest scale there are questions that ask about molesting male children and other questions about molesting female children. Both types of questions are directly related to child molesting, although it is unlikely that the items would be positively correlated since few offenders select both male and female victims.

Due to the method of test construction, individual items on the MSI are often more meaningful than the total score of the scale that comprises the single items. In internally consistent tests, in contrast, the total score would be expected to be more reliable, valid, and meaningful than any individual scale item.

Another serious problem with the MSI is the lack of adequate standardization data. The manual reports information on a few small samples of sexual offenders from inpatient and outpatient settings. The authors acknowledge that these sample sizes are insufficient, although such samples are reasonable to use in the initial stages of scale development. What is most concerning, however, is that most of the psychometric data is supplied on earlier versions of the scales, scales which have since been substantially revised. The manual presents almost no information on the revised scales, except for the means and standard deviations from 40 college students and/or 31 sexual offenders. No psychometric information is presented for some of the scales. Further normative work is clearly required. In particular, the appropriate comparison groups should be used: rather than use convenience samples of college students, efforts should be made to recruit violent and nonviolent criminal controls, and/or men matched on age, education, and socioeconomic status. One criterion for establishing the scales' validities should be their ability to distinguish between sexual offenders and other men.

Nichols and Molinder should be commended for their attention to the problem of denial. The MSI Lie scales, however, are sufficiently ambiguous as to be uninterpretable. Since they are the reverse of the content scales, it is impossible to distinguish between a) offenders who have committed only minimal acts and openly admit to these acts, and b) offenders who have committed numerous and serious acts and deny most of their offenses. Such a distinction has clear implications for treatment responsiveness and risk for reoffending. What is required are minimization/denial scales that are not simply the reverse of the offenders' admitted offense history. Idiographic assessment techniques, which compare a particular offender's own reported offenses to the activities that are documented against him (e.g., police reports, victim statements), are likely to be more clinically useful and defensible than the MSI Lie scales.

The MSI appears useful for predicting outcome in treatment programs that emphasize admission to, and responsibility for, sexual offenses. In agreement with Simkins et al., the MSI appears to be primarily a measure of denial. As such, offenders would be expected to present as less defensive, but more deviant, following treatment. The directly inverse relationship between the denial and the deviance measures on the MSI would hinder its utility for predicting future reoffending. Overall, the strength of the MSI is as a checklist of items to consider when assessing sexual offenders. The lack of attention to internal consistency limits the interpretability of most of the subscales. Nevertheless, the MSI scales are a source of valuable hypotheses to consider in the assessment of sexual offenders. Basic psychometric studies are required before it can be considered to have met acceptable psychometric standards for psychological tests.

Available from:

Nichols and Molinder, 437 Bowes Drive, Tacoma, WA 98466, U.S.A.

Clarke Sex History Questionnaire (SHQ)

Intended use

The SHQ (R. Langevin, Paitich, Russon, Handy & A. Langevin, 1990) is a sexual history survey designed for assessing sexual offenders and other men presenting with sexual problems. It is intended to be used as part of a comprehensive assessment that includes an interview by a professional trained in sexual/psychological assessment.

Description

The SHQ contains 189 direct questions about the frequency of a diverse range of sexual acts. Earlier versions of the scale (Paitich, Langevin, Freeman, Mann & Handy, 1977) had additional questions concerning attraction to, and disgust towards, particular sexual acts; these questions, however, were deleted from the current version because of their high correlations with sex act frequencies. The items in the current version are divided into 14 subscales according to the objects (e.g., male children, adult females) and the activities (e.g., peeping, obscene phone calls) involved. Each item is included in only one subscale. Examples items include "How many TIMES have you kissed or put your mouth between the legs of a girl or woman, since you were 16 years of age," and "How many boys 13 to 15 have you kissed on the lips since you were 21 years of age?" Both of the above questions are answered using the following response options: "(a) none or never, (b) only one, (c) 2-5, (d) 6-10, (e) 11-15, (f) 16-20, (g) 21-25, (h) 26-30, (i) 31-35, (j) 36 or more." The questions can also be scored dichotomously (none/any).

The scales on the SHQ are the following: Female Adult Frequency (20 items), Female Child Frequency (23 items), Female Pubescent Frequency (23 items), Male Adult Frequency (13 items), Male Child Frequency (13 items), Male Pubescent Frequency (13 items), Crossdressing (12 items), Peeping (6 items), Obscene Calls (2 items), Toucheurism & Frottage (6 items), Sadism, Masochism and Rape (16 items), Exhibitionism Frequency (5 items), Exhibiting Behavior (11 items), and Sexual Dysfunction (12 items). As well, it contains questions concerning animal sex (1 item), death (1 item), group sex (2 items), outdoor masturbation (1 item), incest (4 items), fire (2 items), and the current frequency of sexual outlet (4 items).

Most of the scales begin with a general question about whether the respondent has engaged in any activity of a specific type (e.g., sex with male children); if he has, he is to complete the remaining questions in that scale. Otherwise, the respondent is to precede to the next section. The questionnaire takes about 30 minutes to complete. A computer program that administers and scores the SHQ is available (Langevin et al., 1990).

Psychometric evaluation

The SHQ manual reports that the internal consistencies of the scales are generally high (.70 to .98), with the exception of the Toucheurism and Frottage (.69) and Sexual Dysfunction (.51) scales. In general, the internal consistencies of the scales when the items are dichotomously scored tend to be somewhat higher than when the raw frequencies scores are used (Langevin et al., 1990, Table 5).

The authors have found that paedophiles, sexually assaultive men, incest offenders and exhibitionists who admit to their offenses are readily distinguished based on their SHQ responses. The differences are more highly significant using the dichotomous scores than the raw scores for 10 of the 14 scales. The scales are not highly correlated with measures of defensiveness or psychopathology (MMPI L, F and K). Factor analyses of the 14 scales indicate that they are not highly associated with each other. The major exception is that the two exhibitionism scales are correlated with each other and both load on the same factor as the other "courtship disorder" scales - Peeping, Obscene Calls, and Toucheurism & Frottage.

Overall, the manual reports that the standard measures of reliability and validity are acceptable. The evaluation of sex history measures, however, involves some different concerns than the evaluation of sexual attitude measures. Attitudes are latent constructs that are inferred from patterns of responses. The sex acts assessed on the SHQ, in contrast, are sufficiently explicit that responses to single items are interpretable. For example, if an offender indicated on the questionnaire that he had rubbed his penis against two girls 12 or younger (Question 42), the offender's response is important, regardless of his responses to other items.

The merits of sexual history measures hinge upon the clarity of the questions, the ease of administration and scoring, and the range and importance of topics covered; it should strike a balance between comprehensiveness and efficiency. The questions on the SHQ are clear, it is easy to administer, and it is not overly time consuming. Scoring the SHQ, however, is rather cumbersome: the authors recommend against calculating the raw scores by hand and, instead, recommend computer scoring. The dichotomous scores can be calculated by hand, but the authors warn that these scores show only moderate correlations (.65 range) with the scales calculated from raw scores. Their own data, however, show the reliability and validity of the dichotomous scores to be as high, or higher, than the raw scores.

Templeman and Stinnett (1991) used the 1977 version of the SHQ in a study of the sexual preferences and experiences of a sample of "normal" college males. They found that 65% reported some sexually deviant behaviour, although mostly for relatively minor actions, such as secretly trying to see a woman undressing or touching unfamiliar females in crowds. Interestingly, however, three of the sixty respondants reported sexual contact with females under the age of 15 (when they were 20 years old or older) and three reported that they had raped.

The SHQ is the most comprehensive sexual history measure available. It covers a broad range of sexual acts in considerable detail. This detail can be important for comprehensive individual assessments, but it can also be excessive in other contexts (e.g., selection for treatment programming, offender classification). In many contexts, once the offender is identified, for example, as having recurrent sexual relationships with boys, it is not necessary to know the details of the specific sex acts he committed. Consequently, it is not surprising that simplified versions of the SHQ are often used (e.g., Lang, Langevin, Checkley & Pugh, 1987).

Despite its comprehensiveness, there are still areas that it does not cover. One of the most important areas that it does not address is the men's own history of sexual victimization as a child. Although the contribution of childhood sexual abuse to the development of sexual offending has often been overstated (see critical reviews by Garland & Dougher, 1990; Hanson & Slater, 1988), abuse history is, nonetheless, an important consideration in the assessment of sexual offenders (Hanson, 1991; Langevin, Wright & Handy, 1989). Masturbation is another area that is not well covered by the SHQ.

Conclusion

Overall, the SHQ is a promising assessment measure. The scales are clear and the responses to individual items can have considerable clinical utility. It is most useful in identifying and studying the diversity of sexual behaviour in offenders who are willing to admit to their own offenses. Since it is a history measure, it is unlikely to be sensitive to changes in treatment, except for the paradoxical effect of offenders appearing more sexually deviant as they become more self-disclosing. Future studies need to address its ability to predict recidivism and to identify those offenders who could benefit from treatment.

Available from:
Juniper Press, P.O. Box 7205, Oakville, Ontario, Canada, L6J 6L5.

Sexual Interest Card Sort

Intended use

The identification of self-reported sexual preferences in offender and nonoffender populations.

Description and Evaluation

The Abel and Becker Card Sort contains 75 cards on which are printed brief descriptions of various deviant and nondeviant sexual acts (Salter, 1988). The respondent is to place each card into one of seven piles based on how sexually arousing he finds the description. The extreme piles are labelled "extremely sexually arousing" and "extremely sexually repulsive." The card sort can also be easily adapted to a questionnaire format with a 7-point rating scale. The items are sexually explicit. An example item is the following: "A 25 year old man and I are lying side by side naked touching each other all over."

The 75 items are divided evenly over the following 15 content areas (five items each): adult homosexuality, adult heterosexuality, voyeurism, exhibitionism, frottage, homosexual incest, heterosexual incest, homosexual paedophilia, heterosexual paedophilia, transvestism, sadism, masochism, rape, male sexual identity, and female sexual identity. The last two content areas ask direct questions about transsexualism, e.g., "I would like to be a woman."

The Abel and Becker Card Sort is unpublished, although it is available in Salter (1988). Abel, Cunningham-Rathner, Becker and McHugh (1983) reported that some offenders who do not disclose deviant sexual interests in interviews disclose such deviant interests on the card sort. The number who disclose additional deviant interests were not reported in Abel et al.'s (1983) paper, but could be estimated to be about 10%, based on the number of new deviant interests reported and the average number of new deviances for each offender. To date, however, the psychometric properties of their scale have not been formally evaluated.

Laws, Greenbaum, Osborn and Murrin (1989) revised the original Abel and Becker card sort. The Laws et al. measure maintains the same structure as the original measure, but it has been extended to include 130 items. Each content area is addressed by 10 items (increased from 5) and the two sexual identity scales were deleted. Preliminary results with a sample of paedophiles suggested that the internal consistency of the 13 scales was high (.95 to .99 - Cronbach Alphas) (Laws et al., 1989). As well, they found that their card sort was able to distinguish between male object paedophiles and female object paedophiles with the same accuracy as did phallometric testing with slides. All the men in Laws et al.'s (1989) study admitted to their offenses.

Farrenkopt[2] also has developed a sexual interest card sort measure that he has used with juvenile sexual offenders (Templeman & Stinnett, 1991). Farrenkopt's version has items related to rape, exposure, voyeurism, frottage, obscene phone calling, and fetishism. It is also accompanied by a "portfolio" containing photographs of nude males and females of various ages. Templeman and Stinnett (1991) reported that with a sample of "normal" college men, the reliabilities of the Ferrenkopt scales were well within the acceptable range (alpha ranged for .84 to .97). In their study, sexual preferences as measured by the card sort correlated with sexual experiences as measured by the Clarke Sexual History Questionnaire.

Conclusion

The card sort appears to be a promising method of identifying specific sexual interests in deviant offenders. Further research is required, however, before its utility with sexual offenders can be adequately evaluated.

Available in:

The Abel and Becker version is available in Salter (1988). [Salter, A. C., (1988). Treating child sex offenders and victims. Newbury Park, CA: Sage.] Laws et al.'s version is unpublished, but can be obtained from Dr. D. R. Laws, Forensic Psychology, Alberta Hospital Edmonton, Box 307, Edmonton, Alberta, Canada, T5J 2J7.

---

[2] Farrenkopt's measure appears to be unpublished and could not be located in time for this review.

Sex Fantasy Questionnaire

Intended use

This questionnaire was developed by Wilson (1978) to study sexual fantasies and sexual preferences.

Description and evaluation

The scale contains a list of 40 sexual activities. These activities include normal sexual acts (having intercourse with loved partner) and sexual anomalies (e.g., being aroused by watching someone urinate, sex with an animal). Each activity is rated five times according to how frequently it occurs in the following categories: daytime fantasies, fantasies during intercourse or masturbation, dreams while asleep, have done in reality, and would like to do in reality.

The scale is not intended to be an internally consistent scale. The items can, nonetheless, be grouped into the following categories: intimate, exploratory, impersonal, and sado-masochistic. Gosselin and Wilson (1980) provide norms based on samples of normals, fetishists, sadomasochists, and transvestites. The sexual deviants were recruited from various groups for persons with particular sexual interests, such as the Mackintosh Society for fetishists. Although the data was not subject to detailed statistical analyses, the responses for the subjects appear to meaningfully differentiate between the sexual interest groups.

As with all self-report questionnaires, the quality of the responses are determined by the candour of the respondents. This questionnaire, nonetheless, could be useful with sexual offenders who show some willingness to disclose their sexual interests. The activities covered are similar to those of the Clarke Sex History Questionnaire, but none of the responses are self-incriminating. Instead of asking if the subject had sex with children, it asked about sex with a "much younger" person. As well, it is possible for offenders to report a sexual interest while not disclosing that they actually have participated in particular activities.

Available in:

Wilson, G. D. (1978). The secrets of sexual fantasy. London: J. M. Dent & Sons. A simplified version is reproduced in Salter (1988).

Sexual Experiences Survey

Intended use

The Sexual Experiences Survey (SES) was developed to identify "hidden rape victims" and "undetected sexual offenders" in normal samples (Koss & Oros, 1982).

Description

The SES is a short self-report measure that directly asks about sexual victimization and/or sexual assaultive behaviour. A typical item is the following: "Have you ever had sexual intercourse with a man (women) even though you (she) didn't want to because he (you) threatened to end your relationship." The original 12-item scale was later shortened to 10-items (Koss & Gigycz, 1985).

Evaluation

The inherent assumption of the SES is that sexual aggression and victimization are dimensional (i.e., on a continuum) rather than categorical. This approach, consequently, makes the somewhat contentious assumption that there are mainly quantitative differences between sexual "intercourse achieved through verbal coercion" and "intercourse achieved against consent through use of physical force (rape)" (Koss & Oros, 1982, p.455). Therefore, the SES was designed to detect unreported instances of rape and sexual aggression in normal samples (or, more cynically, in convenience samples of college students).

Originally, the SES was a 12 item instrument with a yes-no format regarding sexual intercourse with different types of coercion (one empirically-derived factor). Koss and Oros (1982) reported the results of 3,862 university students' responses. The authors attempted to justify the use of college students by, among other factors, stating that the "work on the instrument was in the development stages" (Koss & Oros, 1982, p.455). Over 8% of the women reported sexual intercourse because of the threat of physical force and 6% reported being raped.

Koss and Gidycz (1985) revised the scale to 10 items and retained the yes-no format. The wording of the items was changed to portray female victimization and male aggression because the authors claimed that "nearly 100% of reported rapes reflect this pairing" (p.422).

The reliability of the 10-item SES appears to be adequate. The internal consistency (Cronbach's Alpha) of the SES using 448 college students was 0.74 for females and 0.89 for males. The authors admitted that there was some "diversity" in female responses (Koss & Gidycz, 1985). The authors also reported that the test-retest reliability of the scale was satisfactory: there was a 93% agreement between the responses of 138 college students on two administrations of the SES, one week apart.

The validity of the SES is debatable. Koss and Gidycz (1985) compared self-report responses on the SES to those of a one-on-one structured interview. The Pearson correlation between female self-report and structured interview victimization responses was 0.73. Twice as many women (16% versus 7.5%) were re-classified from a higher to lower level of sexual victimization on the structured interview. These results suggest that the SES over-estimates sexual victimization in women.

The results with males are even more telling. The SES and structured interview agreement of male sexual aggression was 0.61. Thirty-four percent of the males were re-classified at a lower level of sexual aggression from the structured interviews and only 3% were re-classified as more sexually aggressive. These results emphasize the discrepancy between self-report and structured interview responses of sexual aggression and suggest that the SES may overestimate sexual aggression.

The authors interpreted these results as a form of denial, but it raises serious questions regarding the response biases and demand characteristics of the SES. The subjects who have been used in SES research are primarily young college students (mean age = 21 years in the Koss and Oros, 1982, study) and are likely to be relatively sexually inexperienced. No data on social desirability indices in conjunction with the SES have been reported. Related to this response bias is that of "feminist desirability" where some subjects might give "politically correct" responses that are ultra-sensitive to male sexual "aggression". Similarly, the authors of the SES did not report the wording of instructions given to subjects prior to the administration of the SES. To explain to subjects that they are doing a survey of previously undetected sexual aggression could create demand characteristics that would bias findings. Finally, it is not clear if these reports of sexual aggression and victimization are really "undetected" in the samples. The authors did not mention how many of these subjects had been arrested or treated for sexual offenses.

Despite the shortcomings of the SES, it has been used as a criterion variable for defining sexually aggressive and non-aggressive males in research (Koss, Leonard, Beezley & Oros, 1985; Lisak & Roth, 1988). Lisak and Roth used the SES as a valid measure of male sexual aggression and found that "underlying anger, underlying power, and disinhibition" differentiated sexually aggressive and non-sexually aggressive male undergraduate psychology students based on their responses on the SES.

Conclusions and Recommendations

The advantage of the SES is that it is a short instrument that can be administered to a large number of people in a few minutes. It also includes items that range from mildly to highly coercive. Since it is a direct self-report measure, the adequacy of the responses are based on the offenders' willingness to admit to deviance. It appears highly sensitive to confounding response biases: it appears to overestimate sexual coercion in nonoffender populations, and it is likely to underestimate deviance in applied settings.

Available from:

Koss, M. P., & Oros, C. J. (1982). Sexual Experiences Survey: A research instrument investigating sexual aggression and victimization. Journal of Consulting and Clinical Psychology, 50, 455-457. The 10-item revision is described, but not reproduced, in Koss and Gidycz (1985).

Aggressive Sexual Behavior Inventory (ASBI)

Intended use

The ASBI assesses a history of coercive and violent sexual behaviour in adult males (Mosher, 1988a; Mosher & Anderson, 1986).

Description

The scale directly asks the frequency with which the respondent has engaged in various coercive and aggressive sexual acts. The coerciveness of the behaviours range from mild verbal threats and attempts at manipulation to acts that fit legal definitions of rape. Example items include "I have gotten a woman drunk in order to have sex with her," and "I have forced a woman to have sex with me and some of my pals." Each of the 20 items is rated on a 7-point scale from "never" to "extremely frequent."

Psychometric evaluation

Mosher and Anderson (1986) report a factor analysis of the scale using minimum eigenvalues of 1.0 and varimax orthogonal rotation. They reported that they found six factors with this sample of 125 male undergraduates. Mosher (1988a), however, indicated that this factor analysis was not conducted with the 20-item scale, but with a 33-item scale. Presumably, 13 items were deleted because they did not load on any one factor. Since the factor structure of a scale is highly influenced by the items included, it is unlikely that the six factors identified in the 33-item scale will replicate in the 20-item scale, particularly since the minimum eigenvalue of 1.0 rule is known to overestimate the number of factors (Zwick & Velicer, 1986). Consequently, it is probably best to reserve judgement about the factor structure of the scale until further analyses of the 20-item scale have been conducted.

Mosher (1988a) reported that the internal consistency of the 20-item scale was .94 (Cronbach's Alpha). He also reported a high internal consistency (.87) for a 10-item short form that was administered to a sample of male rock musicians.

Mosher (1988a) reported that the scale correlated positively with his Hypermasculinity scale (Mosher & Sirkin, 1984), and negatively with his Guilt Inventory (Mosher, 1966). Gillis (1991) found that it correlated with the Deitz, Tiemann-Blackwell, Daley and Bentley (1982) Rape Empathy scale and with the Rape scale from the Multiphasic Sex Inventory (Nichols & Molinder, 1984). Gillis (1991) also found that it was not significantly correlated with social desirability (Crowne & Marlowe, 1960), although his sample size in this part of his study was small (N = 16).

Gillis (1991) was the only researcher to use the scale with identified sexual offenders. He found that the ASBI differentiated sexual offenders from nonsexual offenders (Study 2 and Study 3), but did not differentiate sexual offenders from nonsexual violent offenders (Study 3).

The ASBI appears to be a reasonably promising scale for assessing sexual offenders. Like other history measures, its merits are determined not as much by internal consistency, but by clarity of presentation and the importance of the areas covered. It is likely that clinicians working with violent offenders will value the ASBI's descriptive detail concerning coercive sexual behaviours. Considerably more research is required, however, before it is known whether the ASBI is a valid risk indicator. As with all self-report history measures, the adequacy of the responses on the ASBI are primarily determined by the willingness of offenders to admit to deviant sexual practises.

Available in:

Mosher, D. L., & Anderson, R. D. (1986). Macho personality, sexual aggression, and reactions to guided imagery of realistic rap. Journal of Research in Personality, 20, 77-94.
The scale is also reproduced in Mosher (1988a).

Self-Reported Propensity to Rape

Malamuth has designed self-report measures of sexual aggression and propensity to rape. In general, these instruments were developed and used with college students in analogue studies of factors related to rape. Originally the measures were labelled the Likelihood of Raping (LR; Malamuth & Check, 1983; Malamuth, Check & Briere, 1986) and simply consisted of Likert-type ratings of likelihood of raping if there was a guarantee of not getting caught or punished. Later, this instrument was expanded to 14 items and labelled the Attraction to Sexual Aggression Scale (Malamuth, 1989a, 1989b). Malamuth himself has noted that these measures are not designed to identify potential rapists but rather to assess various factors that are associated with some "relative" likelihood of engaging in aggressive sexual behaviour under unspecified conditions (Malamuth, 1989a).

The reliability and validity of these measures in relation to other scales are good. Almost all of the research on these measurements has been done in analogue studies of self-reported propensity for sexual aggression. This propensity has been found to be related to such factors as dominance over women, rape empathy and attitudes towards rape.

Although this analogue research may yield important information into understanding the processes of sexual aggression, it is not clear how these measures could be used in clinical settings, nor does Malamuth claim these measures should be used for such purposes.

Available in:

Malamuth, N. (1989a). The Attraction to Sexual Aggression Scale: Part one. The Journal of Sex Research, 26, 26-49.

Abel's Cognition Scale

Intended use

The scale addresses the rationalizations and cognitive distortions that child molesters may use to justify adult-child sexual relationships. The identified rationalizations can then be addressed in treatment.

Description

The scale contains 28 declarative statements that are rated on a 5-point scale from "Strongly agree" to "Strongly disagree." All the items are worded in the same direction, with agreement indicative of the endorsement of deviant cognitions. Example items include "If a child has sex with an adult, the child will look back at the experience as an adult and see it as a positive experience," and "When a young child asks an adult about sex, it means that she (he) wants to see the adult's sex organs or have sex with the adult."

The scale was originally introduced in Abel et al.'s (1984) treatment manual for child molesters. Due to the popularity of Abel's treatment approach, the scale became widely used by clinicians (Salter, 1988). A French translation of the scale is currently available (I. Pronovost, Centre hospitalier Robert Giffard, Quebec City) It was not until 1989, however, that basic psychometric information on the scale was published (Abel et al., 1989), based on Gore's (1988) doctoral dissertation.

Psychometric evaluation

The Abel et al. (1989) study was based on a group of 240 child molesters, 48 non-child molesting paraphiliacs, and 86 community controls. For the sexual offender groups, the questionnaire was given as part of an extensive clinical evaluation. All the offenders admitted to their sexual deviation. The community controls were matched for age, although they were more well educated than the child molesters, and there were a preponderance of university students in the control group. The factor analysis of the scale showed it to be essentially unidimensional. Thirty-four percent of the variance was accounted for by the first unrotated factor, whereas none of the other factors accounted for more than 4.1%. Although the authors recognize that the scale is unidimensional, they, nonetheless, extract and interpret six factors: "Child-adult sex helps the child," "Children initiate child-adult sex for specific reasons," "Adults initiate child-adult sex for specific reasons," "The child's behaviour shows their desire for child-adult sex," "Adults can predict when child-adult sex will damage the child in the future," and "Child-adult sex is or will be acceptable in society." The internal consistency of the factors ranged from .71 to .84, with the exception of the last factor, which had a Cronbach's Alpha of .59. The one to three week stability of the factors ranged from .64 to .77, with the stability of the overall score being .76.

The child molesters scored more deviantly on all the factors than did the controls, but were not significantly different from the paraphilic group. The paraphilic group was more deviant than the controls on two of the factors (the first and third ones listed above). The only other study that used the scale (Stermac & Segal, 1989), however, did find that the child molesters in their study endorsed more items on the scale than did rapists and community controls. Gore (1988) reported that the child molesters who selected male victims and who had been offending for many years endorsed the most distortions. As well, child molesters endorsed fewer cognitive distortions following treatment. Neither pretreatment nor posttreatment scores, however, were significantly related to sex offense recidivism. Depending on the subscale used, 10% to 30% of the child molesters in Gore's (1988) study endorsed no distortions at pretreatment.

The scale is important because it addresses cognitive factors that are consistent with current theory concerning the maintenance (and perhaps etiology) of child molesting (Abel, Becker & Cunningham-Rathner, 1984; Murphy, 1990; Salter, 1988). Even the authors, however, state that it "must be modified to make it more helpful in assessing child molesters" (Abel et al., 1989, p. 147). In particular, they are concerned that the intent of the scale is highly transparent and that items are all worded in the same direction. They suggest taking the items in the first (rotated) factor and mixing them with other attitude statements that are not associated with child molesting. Initial research with this 6-item factor suggests that it accurately distinguishes between paedophiles and control subjects (Abel, personal communication). Several other researchers are in the process of developing cognition scales modelled after Abel's (Marshall & Hodkinson, 1988, in Hodkinson, 1990; Cortoni, Gordon, Malcolm & Ellerby, 1991). The research on these scales is not sufficiently advanced, however, for these scales to be evaluated at the present time.

Available from:

Dr. G.G. Abel, Behavioral Medicine Institute, Paces Pavilion, Suite 202, 3193 Howell Mill Road NW, Atlanta, Georgia 30327-4101. It is also reproduced in Abel et al. (1989). The French translation is available from Isabelle Pronovost, Centre hospitalier Robert-Giffard, Local M-7500, 2601, de la Canardiere, Beauport (Quebec), G1J 2G3.

Burt's Attitude Scales

Burt's (1980) survey introduced several scales of rape related attitudes. In addition to her widely used Rape Myth Acceptance scale, she also presented scales entitled Adversarial Sexual Beliefs, Acceptance of Interpersonal Violence, Sex Role Stereotyping, Sexual Conservatism, and Sex Role Satisfaction. The latter two scales will not be reviewed since they have attracted little research attention, and there are other measures of these constructs available. Burt's scales are highly intercorrelated, with an average correlation of about .45 (see Burt, 1980; Check & Malamuth, 1983; Demare, Briere & Lips, 1988; Gillis, 1991; Quackenbush, 1989). Nevertheless, each scale will be reviewed separately before returning to a general evaluation of these scales.

Rape Myth Acceptance (RMA)

Intended use

The scale contains items addressing various attitudes that blame rape victims and discredit the credibility of rape reports. The scale was developed for a community survey, but has subsequently been used with sexual offenders.

Description

The RMA scale contains 19 items. The first 13 items are declarative statements to which respondents rate their agreement/disagreement on a 7-point scale, e.g., "If a girl engages in necking or petting and she lets things get out of hand, it is her own fault if her partner forces sex on her." The last six questions rate the credibility of rape reports from particular categories of people, e.g., your best friend, a black woman.

Reliability

Burt (1980) reported a Cronbach's Alpha of .88. Briere, Malamuth and Check (1985), however, found the scale to be multifactorial. Their first two factors concerned 1) disbelief of rape claims (the last question from the scale), and 2) victim responsible for being raped (almost all the other questions). Gillis (1991) reported internal consistency measures of these factors to be .86 and .80, respectively. Briere et al. (1985) also reported a third factor containing two items and a fourth factor containing only a single item. The third and fourth factors are probably best considered noise factors, since they used the eigenvalue-greater-than-one rule, which is known to overestimate the number of factors (Zwick & Velicer, 1986). The authors themselves seem to place little confidence in their factor analysis; although they use the factorially derived subscales in their next publication (Malamuth et al., 1986), Briere returns to using Burt's original scales in later publications (Briere & Runtz, 1989; Demare, Briere, & Lips, 1988). Greendlinger (1985) identified four factors in her factor analysis. Inspection of the first four eigenvalues, however, suggest that the scale is probably unifactorial based on the Scree Test (eigenvalues of 4.8, 2.0, 1.6, 1.4).

Linz, Donnerstein and Penrod (1988) report a Cronbach's Alpha of .71 on a modified seven item version of the scale. In a study designed to change RMA scores, the stabilities of the scores were found to be .77 (pretest to post manipulation) and .69 (pretest to two week followup) (Intons-Peterson, Roskos-Ewoldsen, Thomas, Shirley & Blut, 1989). Higher stabilities would be expected if no intervention was provided. Greendlinger (1985) found the stability to be .74 in a sample of college males (one to two weeks).

Validity

The RMA scale has consistently been associated with liberal, profeminist interpretations of rape vignettes (Burt, 1980; Burt & Albin, 1981; Linz et al., 1988; Quackenbush, 1989). It has also been consistently associated with self-reported likelihood of raping (Demare et al., 1988; Greendlinger & Byrne, 1987; Pryor, 1987; Quackenbush, 1989) and sexual harassment (Pryor, 1987). As well, the RMA scale is moderately correlated with the Attitudes Towards Women scale (Spence, Helmreich & Stapp, 1973), and the Femininity scale (reversed) in Bem's (1981) Sex-Role Inventory (Quackenbush, 1989). Gillis (1991) found that the "victims blaming" factor of the RMA scale correlated with the Hostility Towards Women scale (Check, 1985), the Buss-Durkee Hostility Inventory (Buss & Durkee, 1957), the Rape Empathy Scale (Deitz et al., 1982), and the Cognitive Distortion and Immaturity scale from the MSI (Nichols & Molinder, 1984). Gillis found that the RMA "credibility" factor showed smaller correlations than did the "victim blame," correlations that were only significant for the Hostility Towards Women and Rape Empathy scale. Gillis (1991) also reported that the two RMA factors correlated .24 with each other.

Intons-Peterson et al. (1989) found that the RMA was sensitive to attitude changes produced by exposure to sexually violent films (increase), and education debunking rape myths (decrease). Malamuth et al. (1986) found that it was associated with self-reported arousal to forcing women to do things that they didn't want to do.

Although males scored higher than females on the RMA scale (Ashton, 1982), the scale does not appear to distinguish between sexual offenders and nonoffenders. Burt (1983) found no differences on this scale between incarcerated rapists (in treatment) and the community sample. Overholser and Beck (1986) similarly found no differences between rapists, child molesters, nonsexual offenders, low SES community controls, and minimal dating college males. Gillis (1991) also found no differences between rapists, nonsexual offenders and low SES controls.

Discriminant Validity

Burt (1980) found the RMA to be inversely related to education, SES, and self-esteem, and positively correlated with age. Ashton (1982) found substantial correlations with a Dogmatism scale (.51) (Troldahl & Powell, 1965) and a Trustworthyness scale (-.46) (Wrightman, 1964). It also correlated (.40) with self reported aggression (Greendlinger & Byrne, 1987). Gillis (1991) found that one item from the RMA scale correlated significantly with the Marlowe-Crowne Social Desirability scale, but his analysis lacked power since it only included 16 subjects.

Evaluation

Although further factor analytic studies are required, it appears that the RMA contains at least two factors, a victim blaming factor and a credibility of rape accounts factor. The reliabilities of these subscales are likely to be within an acceptable range. The RMA, particularly the "victim blaming" factor, appears to have moderate construct validity by correlating in the predicted direction with relevant measures. The correlations with other measures are not large enough to question its discriminant validity. Most of the research, however, has been conducted on college males. Its utility with sexual offenders remains questionable. The few studies that have examined the RMA with sexual offenders have found that they respond in a similar manner to other groups.

Acceptance of Interpersonal Violence (AIV)

Intended use

The AIV scale addresses the acceptability of both physical and sexual violence.

Description

Six declarative statements (3 negative; 3 positive) are rated on a 7-point rating scale from "strongly disagree" to "strongly agree." An example item: "A man is never justified in hitting his wife."

Reliability

The internal consistency of this scale was found to be .59 by Burt (1980) and .61 by Malamuth (1986), which is rather weak.

Briere et al.'s (1985) factor analysis divided the scale into three factors. The first factor contained three items concerning women's enjoyment of sexual violence; the second factor contained two items concerning wife assault; and the third factor contained the single item "People today should not use 'an eye for an eye and a tooth for a tooth' as a rule for living." Gillis reported that the internal consistency of the first factor was .74, but for the second factor it was only .47. (Reliability for the third factor could not be estimated since it included only one item.) Greendlinger (1985) found that the first three eigenvalues were 1.8, 1.1, and 1.0, suggesting one weak factor. Overall, the scale appears heterogeneous with little correlations between the items. Marolla and Scully (1986) revised the AIV scale by deleting one item and adding five others, which increased the internal consistency to .72. Their revised scale, however, has not been published and they are the only researchers to have used it.

Validity

The AIV scale showed moderate correlations (.29 to .40) with Nelson's Dominance in Sex scale (Malamuth, 1986), Check's (1985) Hostility Towards Women scale (Malamuth, 1986; Gillis, 1991), Koss and Oros's (1982) Sexual Experiences Survey (Malamuth, 1986), Spence et al.s (1973) Attitudes Towards Women scale (Demare et al., 1988), and Bem's (1981) Sex-Role Inventory (Quackenbush, 1989). Gillis (1991) found that it showed similar correlations with the Buss-Durkee Hostility Inventory (Buss & Durkee, 1957), the Rape Empathy Scale (Deitz et al., 1982), and the Cognitive Distortion and Immaturity scale from the MSI (Nichols & Molinder, 1984). Marolla and Scully (1986) found their revised scale to correlate (.36 to .54) with the Rape Empathy Scale (Feild, 1978), the Attitudes Towards Women scale, and their own version of a hostility towards women scale. Subjects who scored low on this scale were more likely than high scorers to give profeminist interpretations of rape vignettes (Burt & Albin, 1981; Quackenbush, 1989). As well, high scores were weakly associated with self-reported likelihood of raping (Demare et al., 1988; Quackenbush, 1989), sexual harassing (Pryor, 1986) and sexually molesting children (Briere & Runtz, 1989). The scale was also associated with arousal to rape depictions (Malamuth, 1986) and to forcing women to do things that they did not want to do (Malamuth et al., 1986).

All the above studies used community samples, mostly college males. The studies that compared rapists and nonsexual offenders found no differences between them (Gillis, 1991; Marolla & Scully, 1986). Marolla and Scully, however, reported high levels of acceptance of violence in both of these offender groups.

Discriminant Validity

Marolla and Scully's (1986) revised scale correlated weakly (.16), but significantly, with the Marlowe-Crowne Social Desirability scale. Gillis (1991) found no significant relationships with social desirability, although his analysis was based on only 16 subjects. Malamuth (1986) found no significant correlation with Eysenck's (1978) Psychoticism scale.

Evaluation

The scale's internal consistency is sufficiently low as to make it difficult to identify what it is assessing. Considering the scale's low reliability, it shows remarkably strong associations with other relevant measures. There is no evidence, however, that it identifies features more common in sexual offenders than in other populations. The scale is probably addressing relevant constructs, but it needs revision before it meets minimum psychometric standards.

Adversarial Sexual Beliefs (ASB)

Intended use

The scale addresses the extent to which heterosexual relationships are viewed as a "battle," with little possibility of trust between men and women.

Description

The scale contains nine declarative statements (all positive) which are rated on a 7-point scale from "strongly agree" to "strongly disagree." An example item: "In a dating relationship a woman is largely out to take advantage of a man."

Reliability

Burt (1980) reported internal consistency of .80. Briere at al. (1985) found two factors, but since even with orthogonal rotations most items loaded on both factors, the scale is probably best considered unidimensional. Greendlinger (1985) found that only two factors have eigenvalues greater than one (2.5 & 1.3) supporting the view that it contains only one factor. Gillis (1991) reported the internal consistencies of the Briere et al. (1985) factors scales to be .71 (four items) and .38 (three items).

Validity

The ASB has been found to correlate with the Attitudes Towards Women scale (Demare et al., 1988) and with Bem's Sex-Role Inventory (Quackenbush, 1989). Quackenbush (1989) found that it was associated with profeminist interpretations of rape vignettes, as did Burt and Albin (1981), although the relationship in their study was small ($r = .10$, $p < .05$). It has also been associated with self-reported likelihood of raping (Demare et al., 1988; Quackenbush, 1989), and sexual harassment (Pryor, 1987). Malamuth et al. (1986) found that it was associated with self-reported arousal to forcing women to do things that they did not want to do. Gillis (1991) found that one reliable factor from the scale correlated moderately (.41 to .50) with the Hostility Towards Women scale (Check, 1985), the Buss-Durkee Hostility Inventory (Buss & Durkee, 1957), the Rape Empathy Scale (Dietz et al., 1982), and the Cognitive Distortion and Immaturity scale from the MSI (Nichols & Molinder, 1984).

Peterson and Pfost (1989) reported that it increased with exposure to violent rock videos. Their study, however, contained either calculation errors or multiple typographical errors, since the statistics reported do not indicate any departure from chance. They erroneously claimed that an F (3, 140) of 2.56 is significant at $p < .05$; the .05 critical value for F (3, 140) is 2.7. A recalculation from the presented means and standard deviations resulted in marginally lower F value of 2.4.

With the exception of Gillis (1991), all of the above research involved nonoffenders, mostly college males. Gillis (1991) found no differences on the scale between rapists, nonsexual violent offenders, and property offenders.

Evaluation

The scale measures an interesting construct with reasonable internal consistency. It appears unable, however, to distinguish sexual offenders from nonsexual offenders, although there is only one study that used this scale with offenders.

Sex Role Stereotyping

Intended use

The scale addresses conservative versus profeminist attitudes concerning the roles of women.

Description

The scale contains nine declarative statements (7 positive; 2 negative) which are rated on a 7-point scale from "strongly agree" to "strongly disagree." An example item: "A woman should be a virgin when she marries."

Reliability

Burt (1980) reports Cronbach's Alpha to be .80. Malamuth et al. (1986) indicated that it contained two factors: "Stereotyped roles for women in relationships," and "Rules regarding traditional male and female behavior." The details of their factor analysis were not given. Greendlinger (1985) found the eigenvalues of the first three unrotated factors to be 2.1, 1.4, and 1.2, suggesting by the Scree Test that the scale is unidimensional.

Validity

As would be expected, the SRS scale correlates highly (.73) with Spence et al.'s (1973) Attitudes Towards Women scale (Check & Malamuth, 1983). Lisak and Roth (1990) found the SRS scale to correlate moderately (.32 to .56) with Check's (1985) Hostility Towards Women, Nelson's Dominance in Sex, and Lisak and Roth's (1988) own Underlying Power scales. It also was associated with self-reported arousal to rape vignettes (Check & Malamuth, 1983), likelihood of raping (Check & Malamuth, 1983), likelihood of sexual harassing (Pryor, 1987), and arousal to forcing a woman to do something she did not want to do (Malamuth et al., 1986).

Overholser and Beck (1986) found that child molesters scored more conservatively on this measure than did lower class controls or minimal dating college males. Lisak and Roth (1990) found that self-reported rapists - as identified by Koss and Oros' (1982) Sexual Experiences Survey - were more conservative than a control group. Both groups in Lisak and Roth's study were primarily comprised of college students.

Discriminant Validity

Check and Malamuth (1983) found that it had a .32 correlation with a measure of self-reported nonsexual aggressiveness in college students.

Evaluation

The SRS appears to be a reasonable measure of profeminist attitudes. Importantly, it is one of the few questionnaire measures that has distinguished between sexual offenders and nonoffenders. The research is clearly incomplete, and the number of unpublished negative findings are unknown, but the initial results of this scale are promising, particularly considering its short length (9 items). Like the AWS scale, however, its usefulness is likely to be affected by changing social norms concerning attitudes towards women.

Other research

Since the Burt scales are highly intercorrelated, and many are not internally consistent, it is reasonable to consider rearranging the Burt items into entirely new scales, as was done by Koss, Leonard, Beezley and Oros (1985). Koss et al. (1985) identified five factors out of Burt's (1980) original scales, which they titled "Sexual aggression as normal," "Traditional attitudes towards female sexuality," "Rejection of rape myths," "Heterosexual relationships as game playing," and "Rejection of violence." Since the factor analytic methods were not specified, and the items that constituted the new scales were not identified, the adequacy of the new scales is difficult to evaluate. Koss et al. (1985) did report, however, that these scales reliably distinguished between sexually assaultive and nonassualtive college students, as identified by Koss and Oros' (1982) *Sexual Experiences Survey*. No other researchers have used Koss et al.'s new scales, although they are likely to have greater reliability than Burt's original scale.

Conclusions

The Burt scales have strong face validity as measures of rape-related attitudes. Consequently, they may have some utility in sexual offender treatment programs that address conservative, dogmatic sexual views. These scales, however, are most useful as material for discussion. In general, the scales have not distinguished between sexual offenders, nonsexual offenders and nonoffenders, which limits their utility as risk predictors and as indicators of treatment outcome.

Available in:

Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology, 38*, 217-230.

Attitudes Toward Rape Scale (ATR)

Intended use

This instrument was developed by Feild (1978) for the purpose of empirically assessing different groups of individuals' attitudes toward rape in society.

Description

The final form of the Attitudes Toward Rape scale was comprised of 32 positively and negatively worded items on a 6-point Likert scale of agree/disagree. Example items include "A woman should be responsible for preventing her own rape," and "A convicted rapist should be castrated." The scale was accompanied by the Rape Knowledge Test (RKT), a 14-item test in a multiple choice format that was designed to assess individuals' factual knowledge about rape. The RKT mostly concerns demographic information.

Evaluation

The ATR scale was derived from an initial pool of 75 items representing affective, cognitive, and conative components of rape attitudes in the literature. Thirty-seven items were rationally selected from this pool and administered to 400 college students. Item analysis of these responses resulted in the final 32-item format. The RKT was developed in a similar format from a 33-item pool.

Feild (1978) administered these scales, along with the Attitudes Toward Women Scale (AWS; Spence & Helmreich, 1972) to 1,448 subjects that included rapists (n=20), police (n=254), crisis counsellors (n=118), and citizens (n=1,056). Cattell's Scree Test procedure suggested that eight factors should be extracted from the 32-item ATR scale. These factors appeared similar across subgroups as indicated by the coefficient of congruence. The eight factors were interpreted as representing a woman's responsibility in preventing rape, sex and its relation to rape, punishment of a rapist, the role of women in precipitating or causing rape, the mental well-being of rapists, power as motivation for rape, the attractiveness of a woman after her rape, and the expected behaviour of a woman during a rape.

A test-retest study was not conducted so factor reliability is not known from this study. Some of the many significant sex differences included the findings that male citizens tended to indicate to a greater extent that it was a woman's responsibility to prevent rape and that punishment for rape should be harsh. For both males and females, personal knowledge of a raped woman was not significantly correlated with any of the factors. Racial differences included the finding that blacks in comparison to whites felt that women were primarily responsible in rape prevention. There were few significant findings within the rapist group, probably because of the small sample size, and the author did not elaborate, stating that findings could have occurred by chance.

The "Sex as motivation for rape" and "Victim precipitation of rape" factors were significant in a discriminant function that differentiated the groups. The main differences between the groups were between rapists and rape counsellors who appear to lie at opposite ends of the spectrum regarding rape attitudes. Of interest, citizens and especially police officers were more similar to the rapist group in their responses than to the rape counsellor group. It would appear that most differences are a function of gender rather than rapist/non-rapist status. The rapist sample was rather small, however. Finally, the correlations between rape knowledge and attitudes toward rape were weak.

There has been little published research using the ATR scale subsequent to this study. One of the few such studies was by Hamilton and Yee (1990). This study examined the relationship between rape knowledge and propensity to rape (based on self-report) in 276 undergraduates. Instead of the traditional assessment of knowledge of rape demographics or "statistical" information as these authors labelled it, knowledge about the trauma of rape for the victim was assessed. Similar sex differences were found on the ATR scale. Greater knowledge about rape was associated with fewer pro-rape attitudes on the ATR and less self-reported propensity to rape. These results are not surprising since the authors defined knowledge of rape as the aversive and traumatic effects on the victim, both during and after the rape. Correlating this type of knowledge with attitudes toward rape is simply one measure of a construct predicting scores on another measure of the same construct since the measure of rape knowledge was very much a measure of beliefs regarding the impact of rape.

Conclusions and Recommendations

More research on the ATR is needed before any firm conclusions can be drawn regarding its forensic utility. The research to date has found that many of the significant findings relate to gender and traditional stereotypes surrounding rape. It is not clear if sex offenders hold these attitudes more than do males in general, nor whether males with these attitudes will go on to become sex offenders.

Available in:

Feild, H. S. (1978). Attitudes towards rape: A comparative analysis of police, rapists, crisis counsellors and citizens. *Journal of Personality and Social Psychology, 36*, 156-179.

Rape Empathy Scale (RES)

Intended use

The Rape Empathy Scale was developed by Deitz, Tiemann-Blackwell, Daley and Bentley (1982) for the purpose of assessing individuals' empathy toward rape victims and rapists.

Description

The scale contains 20 forced choice alternatives. Example items include "a) In general, I feel that rape is an act that is provoked by the rape victim, b) In general, I feel that rape is an act that is not provoked by the rape victim," and "a) I can understand a wife's humiliation and anger if her husband forced her to have sexual relations with him, b) A husband has every right to determine when sexual relations with his wife occur, even if it means forcing her to have sex with him."

Evaluation

This scale is similar to those developed by Burt (1980) and Feild (1978), and many of the same comments and criticisms apply. Essentially, the findings with this scale are that the more self-reported empathy an individual has for a rape victim, the less likely an individual is to report a propensity to rape. Almost all of the research using this scale has relied on college students.

The item selection was rationally derived from the literature and from pilot studies with college students. The internal consistency was found to be satisfactory (Alphas ranged from .82 to .89). Deitz et al. (1982) developed the scale using college students and potential jurors, but not sex offenders.

Deitz, Littman and Bentley (1984) found that rape empathy was less for "unattractive" female victims. The term unattractive was not operationally defined but rather left to the individuals' interpretation. Most of the effects were significant but not large.

Linz, Donnerstein and Penrod (1988) assessed the effects of exposure to pornographic and violent films on subsequent empathy when viewing a rape trial. One hundred and fifty-six male college students were administered a 12-item version of the RES (Alpha = 0.60). Subjects who viewed violent films subsequently showed less empathy towards rape victims as measured by the RES. The time elapsed between viewing the films and viewing the rape trial was not reported but it appears to have been only a few days. Therefore, the effects may be state dependent. This fact, combined with the use of an analogue sample, suggests the findings may have limited ecological validity.

Gillis (1991) conducted one of the few studies that used the RES with sex offenders. Item 15 from the scale was removed because of a poor item-total correlation, resulting in a revised Alpha of .81. He also found that the RES loaded on a common factor with Burt's (1980) scales. The RES successfully differentiated convicted rapists from convicted property offenders in both of Gillis' studies (1 and 2). However, rapists and nonsexual violent offender controls were not significantly different. Gillis also reported that the offenders had difficulty completing the scale, suggesting further revisions were necessary before it could be clinically useful.

Conclusions and Recommendations

The same comments and criticisms of the Attitudes Toward Rape scale also apply to the RES. Most of the published work on this scale has been with college students and the little research with sex offenders has yielded mixed results. Therefore, the scale's utility in clinical settings is doubtful at this time but more research for this purpose is indicated.

Available in:

Deitz, S. R., Tiemann-Blackwell, K., Daley, P. C., & Bentley, B. J. (1982). Measurement of empathy towards rape victims and rapists. Journal of Personality and Social Psychology, 43, 372-384.

Rape Supportive Attitude Scale (RSAS)

Intended use

The RSAS was developed by Lottes (1988) to "measure beliefs that are hostile to rape victims" (p. 235).

Description and evaluation

Twenty items were selected from an initial pool of 40 items derived from existing rape attitude measures based on high item-total correlations and factor loadings in large samples of college students. Examples of items include "Being roughed up is sexually stimulating to many women," and "A raped woman is a less desirable woman." All items are scored on a disagree/agree Likert format.

Little detailed information is provided on the psychometric development of the scale. Alpha was reported to be .91 and factor analysis suggested that the scale was unidimensional. It correlated in the predicted direction with related measures of callous sexual attitudes and conservative sex role stereotyping (Lottes, 1988). There is no information on this scale other than that provided by Lottes (1988).

The scale appears to have greater internal consistency than other rape attitudes scales. Further research needs to be conducted in order to determine its validity.

Available in:

Lottes, I. L. (1988). Rape Supportive Attitudes Scale. In C. M. Davis, W. L. Yarber, S. L., & Davis (Eds.) Sexuality related measures: A compendium. Lake Mills, IA: Graphic Publishing Co.

Rape Attitude Scale (RAS)

Intended use

The RAS scale was designed to measure attitudes tolerant of rape in offenders and community populations. Adult and adolescent versions are available (Hall, Howard & Boezio, 1986).

Description and evaluation

The RAS contains 14 statements to which the respondents are to rate their agreement/disagreement. A version for university students used a 9-point ratings scale, and an adolescent version used a 5-point ratings scale. Example items include "Most (girls/women) who have sex physically forced on them are easy and loose," and "Most (girls/women) secretly want some guy to use violence to force them to have sex." The scale was constructed by comparing the responses of rapists and low SES community control subjects on several existing rape myths scales (e.g., Burt, 1980; Feild, 1978). Fifteen items that distinguished between these groups were identified - one of these items was subsequently deleted because of a low item-total correlation. The internal consistency of the 14-item scale was .66 for the adolescent sample and .78 for the university sample (Cronbach's Alpha).

Rapists reported more tolerant attitudes towards rape than did community controls on the original 15-item scale. This finding in not particularly surprising, since the items for this scale were selected because they differentiated rapists and nonoffenders. With a sufficiently large item pool, it is possible to construct scales that accurately differentiate randomly assigned groups (even if the items are themselves random numbers). Items that are selected on their ability to distinguish groups need to be replicated on independent samples. Another concern with the RAS is that it did not distinguish between rapists and nonsexual violent offenders. It was unclear whether any attempt was made to identify items that were answered differently by the sexual and nonsexual offenders.

In studies of university students and large random community samples of adolescents (971), Hall et al. (1986) found that their RAS correlated moderately (r = .31) with a revised version of the Gough Socialization Scale (a measure of antisocial behaviour). It also correlated with their Heterosexual Relationship Scale (correlations range from .44 to .54).

Hall et al.'s (1986) scale development strategy seemed reasonable. The scale's low internal consistency, and its inability to distinguish sexual from nonsexual violent offenders, however, limits its utility in the risk assessment of sexual offenders. Additional research is necessary before the merits of this measure can be adequately evaluated.

Available in

Hall, E. R., Howard, J. A., & Boezio, S. L. (1986). Tolerance of rape: A sexist or antisocial attitude. Psychology of Women Quarterly, 10, 101-118.

The Attitudes Towards Women Scale (AWS)

Intended use

The AWS aims to assess conservative versus liberal, profeminist views concerning the roles of women in several areas, including education, occupation, dating, etiquette, and marriage. The scale can be used with offenders and nonoffenders, although it was initially developed for research with nonoffender community samples.

Description

The original version of the AWS contained 55 items (Spence & Helmreich, 1972). Each item of the AWS is a declarative statement to which respondents are to rate their agreement/disagreement based on four alternatives (Agree Strongly, Agree Mildly, Disagree Mildly, and Disagree Strongly). The items are scored from 0 to 3 with high scores indicating liberal, profeminist values. Spence and Helmreich (1972) reported that the original 55 item version was divided into several factors. The factor pattern varied as a function of the age and sex of the respondents; however, since the original scale has been largely replaced by the revised 25 item version, the factor structure of the original scale need not be discussed further. Furthermore, subsequent factor analyses suggested that Spence and Helmreich (1972) extracted too many factors and that both the original and shortened versions were unidimensional (Smith & Bradley, 1980). The revised scale correlated greater than .96 with the original scale (Smith & Bradley, 1980; Spence, Helmreich & Stapp, 1973). Example items from the revised AWS include "Swearing and obscenity are more repulsive in the speech of a woman than of a man," "There should be a strict merit system in job appointment and promotion without regard to sex," and "It is as ridiculous for a woman to run a locomotive as for a man to darn socks." Thirteen of the items are scored in the "Agree" direction and 12 are scored in the "Disagree" direction. A 21 item British version of the scale (AWS-B) is also available (Parry, 1983).

Reliability

Smith and Bradley's (1980) review of the reliability of the AWS suggested that both long and short forms of the scale are internally consistent and stable. They reviewed studies in which the test-retest coefficients were found to be in the .90 range. Internal consistency measures have ranged from .78 to .92 (Feild, 1978; Jean & Reynolds, 1984; Smith & Bradley, 1980). The factor analytic results of both Spence et al. (1973) and Smith and Bradley (1980) indicate that the commonly used short version is unidimensional. Marolla and Scully (1986) reported Cronbach's Alpha to be .77 on a version that contained a "few revisions" intended to limit the middle class bias and to improve readability. Parry (1983) reported Cronbach's Alpha for the AWS-B to be .77 and .85 for lower and middle class women, respectively. Overall, the internal consistency and stability appears within acceptable limits.

Validity

The AWS is found to be associated with a number of other measures that should be associated with liberal, profeminist values. Moderate correlations (.30 to .55) were found with scales such as Burt's (1980) Acceptance of Interpersonal Violence scale (Demare, Briere & Lips, 1988; Marolla & Scully, 1986, using a revised version of the AIV scale), Burt's (1980) Rape Myth and Adversarial Sexual Beliefs scales (Demare et al., 1988), Mosher and Sirkin's (1984) Hypermasculinity Inventory (Archer & Rhodes, 1989), and the Rape Empathy Scale (Deitz et al., 1982). A modified "true/false" version of the AWS similarly correlated with Villemez and Touhey's (1977) Macho Scale, and Nadler and Morrow's (1959) Open Subordination of Women scale (Sattem, et al., 1984). The AWS also showed modest, but statistically significant, correlations (.20 to .30) with Feild's Attitudes Towards Rape scale (Feild, 1978; Marolla & Scully, 1986), Bem's (1974) Sex Role Inventory (Archer & Rhodes, 1989), the Likelihood of Sexually Harassing scale (Pryor, 1987), and a Hostility Towards Women scale developed by Marolla and Scully (1986). (Marolla and Scully's scale should not be confused with Check's HTW scale.) Several studies also found the AWS scale to be associated with profeminist interpretations of rape vignettes, i.e., increased willingness to define the situation as rape, greater sympathy for the rape victims, etc. (Deitz et al., 1982; Marolla & Scully, 1986; Muehlenhard, Friedman & Thomas, 1985).

Demare et al. (1988) found the AWS to be associated with self-reported likelihood of raping. Another study found that male students who reported actually having forced nonstranger females to commit sex acts also showed the most conservative attitudes on the AWS (Koss, Leonard, Beezley & Oros, 1985). None of the respondents in Koss et al.'s study reported that they were charged for their actions, although the actions they reported would fit legal definitions of rape.

Despite the promising correlations with other questionnaire measures, the AWS has generally not been effective in differentiating groups of identified sexual offenders from nonoffenders. Several studies found that the AWS failed to discriminate between rapists, child molesters, nonsexual offenders, and low SES community controls (Marolla & Scully, 1986; Sattam et al., 1984; Segal & Marshall, 1985; Segal & Stermac, 1984). Scott and Tetreault (1987) provided the only study in which rapist scored more conservatively than violent criminals and SES matched male community controls. The average values of all their groups, however, were high. The rapists had an average score of 57.2, which was .6 standard deviations higher than the highest means reported by Spence et al. in 1973 (50.3 for the university educated females). Scott and Tetreault's community control group almost provided the most profeminist responses possible (average of 73.9 with a maximum score of 75). Since most of the offenders were serving long sentences, one interpretation of these results is that the offenders were less exposed than the community controls to the changes in attitudes towards women that have permeated popular culture. Alford and Brown (1985) similarly found that a group of rapists scored significantly higher than the norms for males reported by Spence and Helmreich in 1972.

Discriminant Validity

The AWS does not appear to be correlated with standard measures of social desirability, such as the Marlowe-Crowne Scale (Marolla & Scully, 1986). Nevertheless, when given directions to fake liberal, fake conservative, or respond as they normally would, the normal responses of both female and male students were not significantly different from their responses in the fake liberal condition (Jean & Reynolds, 1984). The responses in the fake conservative condition, however, were significantly more conventional than their responses in the other two conditions. One possible interpretation is that the subjects' normal response set was an attempt to identify the most liberal responses.

Another potential confounding factor is that the AWS appears to be related to social class. Segal and Marshall (1985), in the only study that examined social status differences, found that high SES respondents gave the most liberal, profeminist attitudes. Segal and Marshall also found that the AWS scale was positively associated with assertiveness, but this correlation was probably mediated by the relationship of both assertiveness and the AWS to social class.

Overall summary and evaluation

The AWS appears to be an internally consistent scale with reasonable construct validity. It has shown the expected pattern of correlations with similar questionnaire measures, and, at least with nonoffender populations, appears to be a reasonable measure of liberal, profeminist attitudes. Its utility with sexual offenders, however, appears to be limited. The research to date suggests that sexual offenders typically do not differ from nonoffenders on this scale. This lack of differentiation may be attributed to at least two possible problems: 1) the attitude assessed by the scale may have little direct relationship to sexual offending, and/or 2) the scale may be unable to accurately detect existing attitudinal differences in the sexual offenders (e.g., the offenders may be motivated to present socially acceptable responses). A further concern is that many of the items are too strongly worded considering the current range of attitudes towards women that are acceptable in western culture; consequently, as attitudes become increasingly liberal, the scale loses sensitivity to existing variations in profeminist attitudes.

Available in:

The short form is reproduced in
Spence, J., Helmreich, R. L., & Stapp, J. (1973). A short version of the Attitudes Towards Women Scale (AWS). Bulletin of the Psychonomic Society, 2, 219-220.

Bias In Attitudes Survey (BIAS)

Intended use

The BIAS was developed by Jean and Reynolds (1980) for the purpose of developing a sex bias measure that was more sensitive than existing measures.

Description and evaluation

It contains 35 items rated on a disagree/agree Likert scale. Examples of items are "The equal rights amendment is unnecessary," and "Little girls are more easily frightened than little boys." Its internal consistency was within acceptable limits. Jean and Reynolds (1980) found in a sample of 354 undergraduates that males, younger students, and divorced students tended to score the highest on the scale. In a subsequent study, Jean and Reynolds (1984) found that it was more difficult to fake than the AWS. Males, in particular, showed little difference in their scores when asked to fake liberal or conservative. Females, however, were able to present conservative responses on request.

Overall, the BIAS appears to be a more sensitive measure than the AWS for assessing liberal versus conservative sexual role values. It is questionable, however, if this construct has utility with sexual offenders.

Available in:

Jean, P. J., & Reynolds, C. R. (1980). Development of the Bias in Attitude Survey: A sex-role questionnaire. The Journal of Psychology, 104, 269-277.

Index of Sex-Role Orientation

This scale was developed by Dreyer, Woods & James (1981) because the authors felt that sex-role research had been hampered by the lack of valid and reliable research tools. The scale is a 16-item questionnaire designed for women. Example items include "On the job, men should not refuse to work for a woman supervisor," and "Although women hold many important jobs, their proper place is in the home." The authors report a corrected split-half correlation of .92. The one-month stability coefficient for 32 female students was .62, which is moderate to low. Members of a feminist organization significantly differed from members of church-sponsored women's groups. It was also highly associated with education. Like the BIAS, this scale appears to be a more sensitive measure of sexual role attitudes than the AWS.

Available in:

Dreyer, N. A., Woods, N. F., & James, S. A. (1981). ISRO: A scale to measure sex-role orientation. Sex Roles, 7, 173-182.

Sexism Scale

This scale was developed by Rombough and Ventimiglia (1981) as a measure of sex-role attitudes. A pool of 197 items that were derived from previous research and literature were administered to a "diverse sample" of 154 respondents. A factor analysis suggested that the scale was unidimensional. This analysis, however, should be considered cautiously since the item to subject ratio was low. The authors rationally divide the scale into two sets of items: internal versus external division of labour, and sex differences. Examples of items include "Men make better engineers than women," and "Women should not have the same sexual freedom as men." The final scale was composed of 20 items scored in a Likert agree/disagree format. Alpha reliabilities ranged from .73 to .88. Little validity research has been conducted. The scale may have promise in sexual role research; however, its psychometric development appears weak and there is little to recommend it over the other available sex role measures.

Available in:

Rombough, S., & Ventimiglia, J. C. (1981). Sexism: A tri-dimensional phenomenon. Sex Roles, 7, 747-755.

Attitudinal Inventory

This scale is a 21-item self-report measure of feminist versus traditional sex roles (Smith & Self, 1981). Items include "Women should feel free to ask men for dates," and "Husbands should be the head of their households." The items are rated on a Likert agree/disagree scale.

The scale was developed for use in a study of feminist and traditionalist values amongst female college students. The different groups of women were determined based on their agreement to the statement "I consider myself a feminist." Most of the correlations between the scale items and the different value groups were quite low (.06 to .48). The scale appears to have limited utility, although the authors concluded that feminists display a more "consistent sex-role ideology" because they had more inter-item consistency on this scale.

Available in:

Smith, M. D., & Self, G. D. (1981). Feminists and traditionalists. Sex Roles, 7, 183-188.

Hypermasculinity

Intended use

The Hypermasculinity scale is intended to assess "macho" personality characteristics in adult males.

Description

The scale asks respondents to select from two alternatives - one representing "macho" values whereas the other represents "nonsexist" values. Example items include "I like fast cars and fast women; I like dependable cars and faithful women." The scale is divided into three subscales: a) Calloused sex attitudes towards women; b) violence as manly; and c) danger as exciting. Each subscale contains 10 items.

The scale is based on earlier work on calloused sexual attitudes (Mosher, 1971a, 1971b). The current revision (Mosher & Sirkin, 1984) was developed by selecting from 90 items those items that showed the highest item-total correlations with the three rationally determined subscales. The ten best items from each subscale were then combined into the 30 item Hypermasculinity scale.

Psychometric evaluation

Mosher and Sirkin (1984) reported that the combined 30 item scale contains a single factor. They were unclear about the criteria they used for determining the number of factors; nonetheless, they did report that the first unrotated factor accounted for 24% of the variance, whereas none of the other factors accounted for more than 7%. Based on the Scree Test (Cattell, 1966), the above pattern of eigenvalues suggests that the scale contains a single underlying factor. The internal consistency of the total scale was .89 (Cronbach's Alpha); the internal consistencies of the shorter subscales were predictably lower: Calloused Sex (.79), Danger (.71), and Violence (.79). Mosher and Sirkin's (1984) factor analysis did not find the scales to represent distinct factors. The subscales intercorrelated quite highly in both Archer and Rhodes (1989) (.36 to .51) and Mosher and Sirkin's (1984) samples (.58 to .63).

The Hypermasculinity scale has shown convergent validity by correlating with other relevant scales, such as Spence et al.'s (1983) Attitudes Towards Women scale (Archer & Rhodes, 1989), Bem's (1974) Sex-Role Inventory (Archer & Rhodes, 1989), and Mosher and Anderson's (1986) Aggressive Sexual Behavior Inventory. Mosher and Sirkin (1984) also reported that it was associated with self-reported alcohol and drug use, dangerous driving and juvenile delinquency. It correlated only slightly with the Social

Desirability Scale from the Jackson's Personality Research Form (Jackson, 1974).

The construct assessed by the Hypermasculinity scale appears to be more relevant for sexual offenders than many of the other sex role measures considered in this review. In particular, its emphasis on violence and risk taking are factors that are known to be associated with criminality (e.g., Hare et al., 1990). As well, the scale appears to be internally consistent and not overly contaminated with social desirability. The research on the Hypermasculinity scale, however, has exclusively used male undergraduates. Consequently, it is difficult to form any conclusions about the usefulness of this scale. Although the research has yet to be conducted, the scale is likely able to differentiate offenders from nonoffenders, but unlikely to differentiate sexual offenders from other types of criminals.

Available in:

Mosher, D. L., & Sirkin, M. (1984). Measuring a macho personality constellation. Journal of Research in Personality, 18, 150-163.

Heterosexual Relationships Scale (HRS)

Intended use

The HRS was designed to assess, in both adults and adolescents, sexist attitudes specific to heterosexual relationships.

Description and evaluation

Rather than provide a global assessment of sexist attitudes, the HRS sought to assess male domination specific to heterosexual interactions. An unknown number of items were selected on a theoretical basis and screened for face validity by a group of 20 university students. Some of these items were dropped from the final versions due to low item-total correlations, resulting in a 12-item adolescent version and a 9-item version for university students. The internal consistency of both of these scales was somewhat low, .64 (Cronbach's Alpha). The scale correlated with a revised version of the Gough Socialization Scale, Hall et al.'s (1986) own Rape Acceptance scale, and with profeminist interpretations of rape vignettes (Howard, 1984). Males endorsed more sexist attitudes than did the females in both the adolescent and university student groups.

Although the current version of the HRS appears heterogeneous and has not been adequately researched, the intent of the scale seems more relevant to sexual offenders than many of the other omnibus sexism scales (e.g., Attitudes Towards Women, BIAS). If sexual offenders are going to have more sexist attitudes than nonoffenders, such attitudes should be most clearly evident in their views concerning sexual relationships. In contrast, the theoretical connection is indirect between rape and sexist views concerning employment equity or manners.

Available in

Hall, E. R., Howard, J. A., & Boezio. (1986). Tolerance of rape: A sexist or antisocial attitude. Psychology of Women Quarterly, 10, 101-118.

Derogatis Sexual Functioning Inventory (DSFI)

Intended use

The DSFI was designed to provide a broad assessment of individuals presenting with sexual dysfunction (Derogatis & Melisaratos, 1979). The scales address sexual attitudes, sex knowledge, specific sexual problems, and general measures of subjective wellbeing.

Description and evaluation

The DSFI inventory is a questionnaire measure that takes about 30 to 45 minutes to complete. There are 10 subscales: Information, Experience, Drive, Attitudes, Psychological Symptoms, Affects, Gender Role Definitions, Fantasy, Body Image, and Sexual Satisfaction (Derogatis, 1980). Each one will be briefly described. The Information subscale is a 26-item true/false test of basic sexual information. It has strong face validity, but only moderate internal consistency (.56) and 14 day stability (.61). The internal consistency and stability of all the other scales appear within acceptable limits (exceptions noted). The Experience subscale is a list of common sexual acts and asks the respondent if he or she has engaged in these activities (e.g., intercourse-female superior position, mutual oral stimulation of genitals). It resembles the scales of Zuckerman (1973) and Bentler (1968). Like these other scales, it is highly reliable.

The Sex Drive scale attempts to assess sex drive through the frequency of sexual activity. It contains seven items and is moderately reliable. The Attitude scale addresses Liberal and Conservative sexual attitudes. Derogatis considers these attitudes to be orthogonal, although the scale score is constructed by subtracting the Liberal subscore from the Conservative subscore.

The Symptoms and Affects scales are general measures of subjective distress and psychopathology. The Gender Role scale involves self-ratings on 30 trait adjectives. These adjective correspond to tradition stereotypes concerning personality characteristics of males (decisive, rational, mechanical) and females (sympathetic, flirtatious, fashionable). Since these scales are only indirectly related to sexuality and sexual offending, they need not be considered further.

The Fantasy scale asks respondents to indicate which of 20 fantasies they have had, either as daydreams or asleep. These fantasies contain both normal and deviant sexual practises. This scale is a simpler fantasy scale than Wilson's (1978). The comparative validity of these scales has not been assessed.

The Body Image scale addresses how pleased the respondents are with their physical appearance. It contains 10 Likert ratings for both sexes, and then five separate questions for each sex. An example item, "There are parts of my body I don't like at all." The internal consistency of this scale is only marginal (.58). The Satisfaction scale contains 10 true/false items concerning satisfaction with one's sexual life, e.g., "Often, I worry about my sexual performance."

Considerable research has been conducted with the DSFI as a measure of sexual dysfunction in clinical populations, and it appears to be a reliable and valid measure for this purpose (see Derogatis, 1980). Little research has used this scale with sexual offenders. The studies that have, however, found that sexual offenders report high levels of dissatisfaction with their current sexual life (Dwyer & Amberson, 1989; Pawlak, Boulet & Bradford, 1991). The significance of this finding is difficult to assess. It is easy to imagine that sexual offenders will be unhappy with their sex life once they have been apprehended. It is not clear, however, that sexual dysfunction in any way contributed to their offending.

Available from:

Clinical Psychometric Research, Inc., P.O. Box 619, Riderwood, Maryland, 21139. Toll free number: 1-800-245-0277.

Negative Attitudes Towards Masturbation

Intended use

The scale measures masturbation guilt and negative attitudes towards masturbation. It is was designed for educated males and females (Mosher, 1988b).

Description

The 30 items of this scale are declarative statements that are rated on a 5-point scale from "Not at all true" to "Strongly true for you." Example items include "I feel guilty about masturbating," and "Excessive masturbation leads to mental dullness and fatigue." Twenty of the items are scored in the positive direction and 10 are scored in the reverse direction. The scale is designed to be unidimensional.

Psychometric evaluation

Abramson and Mosher (1975) report that a factor analysis "using an orthogonal varimax procedure" yielded three factors. They appear to confuse the extraction and rotational procedures. Nevertheless, they report that the variance accounted for by the first three factors were 32%, 7% and 6%. Presumably no other factors have eigenvalues greater than one (3.3% of the variance). The Scree Test (Cattell, 1966) suggests one factor and possibly three. The factor loadings for the three factor orthogonal rotation solution, however, do not identify clear factors. Twenty percent of the items load greater than .30 on two factors, and 50% load greater than .25 on two or more factors. Furthermore, the first factor is almost entirely composed of the 10 items that are scored in the reserve direction, which suggests that it is a spurious response-bias factor. Until further factor analytic studies are conducted, it is probably best to consider it a one factor test.

Abramson and Mosher (1975) reported a corrected split-half internal consistency estimate of .75. They fail to identify how they "split" the scale. Mosher (1988b) reports that in another sample of college women the internal consistency was estimated at .94 (Cronbach Alpha). Mosher (1988b) also reviews several studies suggesting that low scores on the scale are associated with decreased sex guilt, increased sexual activity and increased sexual responsiveness in samples of college males and females.

Cotten-Hustan (1983) administered the scale to a small group of sexual offenders as part of a study on the effects of sexual education. For university students, the sex education program decreased negative attitudes towards masturbation. The results for the sexual offender groups were not consistent and difficult to interpret since so few were in the program (3 child molesters; 8 rapists) and control group (5 child molesters; 7 rapists). Although the overall differences between the offender and nonoffender groups would not be expected to be statistically significant with this small sample, it is noteworthy that the sexual offenders tended to report less negative attitudes towards masturbation than did the college students.

The masturbation attitudes and practises of sexual offenders are important concerns in the assessment and treatment of sexual offenders. This scale, however, is not adequate for this task. The assumption that it is healthier to not feel guilty about masturbation seems reasonable for nondeviant samples, but is not appropriate for offenders who masturbate to deviant sexual fantasies (e.g., children, violence). The only study that has examined the relationship between "unhealthy attitudes towards sex" and sex offense recidivism found, contrary to expectation, that adolescent sex offenders who had the most healthy attitudes were the most likely to reoffend (Smith & Monastersky, 1986). Unhealthy attitudes towards sex were not explicitly defined in this study, but it is reasonable to assume that it included guilt concerning sexuality.

A masturbation measure for sexual offenders needs to be developed that addresses not only guilt, but also masturbation frequency, masturbation fantasies, and the circumstances that lead up to masturbation. For example, it would be important to know if an offender masturbates to violent fantasies once a week after watching pornography, or if he masturbates to fantasies about children after his daily visit to the video arcade.

Available in:

Abramson, P., & Mosher, D. L. (1985). Development of a measure of negative attitudes towards masturbation. Journal of Consulting and Clinical Psychology, 43, 485-490.

Hostility Towards Women Scale (HTW)

Intended use

The scale addresses resentment and anger towards women (Check, 1985). The scale and construct are assumed to be trait in nature. It was intended to be used when participants' responses are anonymous.

Description and evaluation

The scale is composed of 30 agree-disagree items. The items are seen to reflect resentment of women (e.g., "I get upset by even slight criticism from a woman"), suspicion of women (e.g., "It is safer not to trust women") and, to a lesser extent, guilt over women and indirect hostility.

Check initially generated a 188 item pool derived from the literature on hostility against women. Thirty items were selected based on item-total correlations using 136 male college students.

Check (1985) reported high one-week stability (0.83) and internal consistency (0.80). Gillis (1991) reported the internal consistency to be .85 (KR-20) in a mixed sample of sexual and nonsexual offenders. Check's factor analysis indicated that the scale is a unidimensional measure, although the single factor only accounted for 25% of the scale variance (Check, 1985). The scale did not correlate with social desirability when given under anonymous conditions to college students; however, the associations to social desirability in offender populations have been very high: -.75 in Gillis (1991), and -.57 to -.60 in Hall (1989).

Check (1985) conducted a series of studies with college males to assess the validity of the scale. The HTW was found to significantly correlate with self-report measures of aggressive attitudes and behaviour, but this was less true for laboratory-assessed behaviours. The HTW did not correlate with aggressive behaviour towards a female target but did towards a male target in a laboratory setting. Check suggested that the scale did not correlate as well with overt aggression or "real world" aggression against women (e.g., rape).

There has been little research on the HTW in forensic settings. Gillis (1991) found that it did not distinguish between rapists, violent offenders, or nonviolent property offenders.

Conclusions and Recommendations

The HTW is a reliable measure but the validity results are mixed. It correlates with other "attitudes" measures but generally not with overt sexual aggression. Its high correlations with social desirability hinder its utility in applied settings.

Available in:

Check, J. V. P. (1985). The Hostility Towards Women Scale (Doctoral dissertation, University of Manitoba, 1984). Dissertation Abstracts International, 45(12). It is also available directly from J.V.P. Check, Psychology Department, York University, 4700 Keele Street, North York, Ontario, M3J 1P3.

Personal Constructs Assessment

Howells (1979) and Horley (1988) have applied Kelly's (1955) Role Construct Repertory Grid Technique (REP) to the assessment of "cognitions" in sex offenders. This approach involves a content analysis of how the offenders identify similarities and differences between people in particular roles (e.g., best friend, abused child).

Howells (1979) used this type of technique in a content analysis of REP grid data to compare the personal constructs of child molesters and non-sex offender prisoners. Few differences emerged but Howells did find that child molesters, compared to property offenders, viewed children as passive and undemanding.

Horley (1988) attempted to replicate Howells' study using nine incarcerated child sexual abusers and 10 incarcerated non-sexual offenders. Two forms of the REP grid were used, one with 10 male elements and the second with 10 female elements. Ten construct categories were generated from the responses including sexuality and morality. The male prisoners used significantly more sexual and appearance constructs when categorizing males compared to females. No between group differences were observed and Howells' (1979) results were not replicated.

Conclusions and Recommendations

The results of studies using Kelly's sorting techniques have not been impressive in differentiating sex offenders from other incarcerated males. This lack of findings may be largely due to the small sample sizes employed in the studies. It is also a relatively time consuming procedure that requires special training to administer and interpret. Although it has the potential to avoid some of the response biases associated with most questionnaire measures, the research has yet to support its utility with sexual offenders.

Available from:

General forms of the REP grid are described in Kelly (1955). [Kelly, G. A. (1955). *The psychology of personal constructs* (Vols. 1 & 2). New York: Norton.] Researchers, however, would want to adapt the elements to make it more appropriate for sexual offenders.

The Minnesota Multiphasic Personality Inventory (MMPI)

The MMPI is one of the most commonly used self-report measures of psychopathology in clinical practice.[3] It should not be surprising, therefore, that a significant amount of research has investigated the utility of the MMPI specifically with child molesters. Unfortunately, research on the MMPI, or variations of it, for this purpose has not been very encouraging other than demonstrating overall differences between sex offender and normal samples. The majority of research in this area can be divided into four areas: 1) the development of specific sex offender scales using MMPI items; 2) overall MMPI scale differences between sex offenders and normals; 3) MMPI profiles associated with specific types of sex offenders; and 4) the utility of two-point MMPI profiles in classifying sex offenders.

Sex Offender Scales Derived from MMPI Items

Probably the most well known scale of this type is the Thorne Sex Inventory (Thorne, 1966b). Although the sex inventory contains many additional items, many of the items were originally adapted from the MMPI. The Pedophilia Scale (Toobert, Bartelme & Jones, 1959) and the Sexual Deviation Scale (Marsh, Hilliard & Liechti, 1955) were also both empirically derived from MMPI item analyses. Toobert et al. (1959) analyzed the MMPI item responses of 120 male pedophiles (both heterosexual and homosexual) and 160 male prisoners with non-sexual offenses as a control group. The Psychopathic Deviance and Depression scales were elevated in both groups, which is a common finding, although neither group had scales elevated over *T* = 70. The authors reported that the Masculinity-Femininity and Paranoid scales "significantly separate the groups in a statistical sense" (p.273). Pedophile responses were compared to those of 139 normal Minnesota males and it was found that there were "72 items which differentiated the groups at a significance level of 5 per cent or better" (p.275). The pedophiles and prisoners were then compared on these 72 items and this analysis resulted in a final pool of 24 items. There was no mention of a Bonferroni-type connection for the number of comparisons relative to the sample size. The Pedophilia scale consisted of sexual as well as religious items but was not reliably cross-validated with another sample of pedophiles in the study.

Similarly, Marsh et al. (1955) conducted a MMPI item analysis using 100 sex offender males and 100 college students (gender distribution was not reported). There were 100 items that differentiated the groups at a significance level of 0.01 or better. The 100-item Sexual Deviation Scale was then cross-validated with groups of 200 and 138 sex offenders and 145 and 172 students. This type of cross-validation is limited, however, by the fact that the groups were not really _independent_ samples of subjects. Furthermore, this type of study treats sex offenders as a homogenous population, neglecting important differences between types of offenders. The methodological weakness inherent in these two studies no doubt is largely responsible for the lack of clinical and research interest in these scales. Recently, Langevin, Wright & Handy (1990a, 1990b) have reported Cronbach Alphas of 0.62 and .57 for the Pedophilia scale and 0.88 and .84 for the Sexual Deviation scale, using samples of 479 and 157 sex offenders. In their studies, the Pedophilia and Sexual Deviation scales correctly differentiated sex offenders and controls 68.1% and 70.6%, respectively (50.0% could be correctly classified by chance alone) (Langevin et al., 1990a). Langevin et al. (1990b) reports additional psychometric information on a variety of sexual pathology scales, but concludes that "only the Sexual Deviation Scale had sufficient internal consistency to be considered further for clinical use" (p. 477).

---

[3] Sex offender research on other personality scales such as with the Eysenck Personality Inventory (Malamuth, 1986), Edwards Personal Preference Schedule (Fisher & Howell, 1970) and California Personality Inventory (Freund, Langevin, Laws & Gerber, 1974) is sparse and the findings are generally unremarkable.

Sex Offenders and Normals' Responses on the MMPI

The most common finding in this line of research is that sex offenders have elevated scores on the Psychopathic Deviance (Pd) scale of the MMPI compared to normals. This is true for several types of sex offenders, particularly rapists (Kalichman, Craig, et al., 1989; Kalichman, Szymankowski, et al., 1989). Elevated Pd scores in sex offenders have also been reported in cross-cultural studies (e.g. Velasquez, Callahan & Carrillo, 1989).

Fewer differences have been found among specific types of sex offenders. For example, Scott and Stone (1986) found that although both natural father and stepfather incest perpetrators were somewhat antisocial (Pd scale), their clinical scores were not pathological ($T < 70$), and the two groups were very much alike in their responses. Similarly, Panton (1979) found that 35 incest males and 28 non-incest male pedophiles had very similar MMPI profiles with the incest males scoring higher on the Social Introversion (Si) scales. As is common in most of these studies, the highest scale was the Pd scale followed by a measure of emotional distress, in this case the Depression (D) scale. Finally, Paitich et al. (1977) reported that the F-scale from the MMPI validity scales (general maladjustment or "faking bad") correlated 0.22 with the frequency of heterosexual pedophiliac outlet and 0.23 with the desire for pedophiliac outlet.

Kalichman (1991) similarly found that the MMPI did not discriminate between homosexual and heterosexual child molesters.

He did find, however, that the child molesters showed greater levels of disturbance on many of the MMPI scales when compared to rapists, particularly on the Pa and Sc scales. Offenders against adolescent females tended to show disturbance that was intermediate between the responses of the child molesters and rapists. As in other studies, most of the profiles of the offenders were within the normal range.

MMPI profiles and types of sex offenders

Langevin and his colleagues (Langevin, Paitich, Freeman, Mann & Handy, 1978; Langevin et al., 1990a, 1990b) have presented the results of several analyses of MMPI responses from their sex offender data base. Langevin et al. (1978) used discriminant function procedures to classify groups of 29 heterosexual pedophiles, 22 homosexual pedophiles, 27 incest perpetrators, 217 sex offenders with multiple types of offenses and 54 non-psychiatric heterosexuals. The Masculinity-Femininity (Mf) scale was somewhat useful for differentiating heterosexual versus homosexual offenders (71% accuracy) although this scale did not differentiate heterosexual pedophiles and homosexual pedophiles. Pedophilia, in general, was not associated with femininity. The authors reported that heterosexual pedophiles were more conservative than homosexual pedophiles but otherwise there were few meaningful differences. In general, pedophiles were more emotionally distressed than controls, but as the authors pointed out, this common finding is likely the result of stigma and ostracism. Social introversion was not a good discriminator of incest, pedophilia, and exhibitionism, although pedophiles scored significantly higher on this scale compared to other offenders and controls.

Using almost the same data set (N=479), Langevin et al. (1990a) compared types of sex offenders on 125 variations of the MMPI ("research scales"). They reported both the traditional MMPI norms and the more liberal Colligan-adjusted norms (Colligan, Osborne, Swenson & Offord, 1984). Only 39 scales (31%) had internal consistency alphas of 0.80 or better. In a subsequent sample of 157 sexual offenders, 41 of the 125 scales examined had reliabilities greater than .80 (Langevin et al., 1990b). Most of the scales derived from the MMPI, such as Aggravated Sex, Homosexuality and Romantic Interest, had low reliability figures and were not good discriminators of sex offenders versus normals. The Pedophilia scale correlated -.31 with education level ($p < .001$). The Sexual Deviation scale correlated -.35 with education, -.33 with Raven IQ, and -.50 with age (all significantly at $p < .001$). Several other scales also significantly differentiated sex offenders from controls. These scales included alcohol abuse scales, personality scales such as Dominance, Dependency, and Psychopathic Manipulation, violence scales to a lesser extent, and even brain damage scales. The relative ability of each significant scale to differentiate among sex offender types was not examined. Rather, the authors only reported percentages of offender types with elevated scale scores ($t < 70$).

The most common elevations for heterosexual pedophiles were on Pd (62.1%), D (55.2%), Sc (55.2%) and Psychasthenia (51.7%) using the old norms and on Sc (31.0%), Pd (27.6%), D (27.6%), and F (24.1%) using the Colligan norms. The most common elevations for homosexual pedophiles were on Mf (52.4%), Pd (47.6%) and D (47.6%) using the old norms and on Sc (28.6%), Pd (28.6%), Pa (23.8%) and D (23.8%) using the Colligan norms. The significance of between-groups differences were not reported. The number of comparisons and discriminant function analyses done in these studies raises the likely possibility that many findings could be due to chance alone. The significance of many of the discriminants was weak and therefore the clinical, rather than statistical, significance may be negligible.

Two-point Codes in Sex Offenders

Because of the apparent lack of strong findings using MMPI scales, some researchers have studied two-point codes or patterns of elevation in various types of sex offenders. Armentrout and Hauer (1978) studied two-point profiles in 21 male pedophiles and 13 rapists. Rapists tended to have a Sc-Pd profile; pedophiles had a less elevated Pd-Sc profile. There were, however, no significant differences between the two groups of sex offenders.

In what is probably the largest study of this type, Erickson, Luxenberg, Walbek and Seely (1987) examined the valid profiles of 403 sex offenders with various offenses. The most common two-point codes in pedophiles (N=207) were Pd-Sc/Sc-Pd (13.0%), Pd-D/D-Pd (12.6%) and Pd-Mf/Mf-Pd (10.1%). Pedophiles were significantly more likely ($p < .01$) to have Pd-D/D-Pd codes compared to offenders against adult women and these latter offenders were significantly more likely ($p < .025$) to have Pd-Hypomania/Hypomania-Pd codes compared to pedophiles. The Pd scale was the most common element of the two-point codes for both types of sex offender. In addition, there was a lot of alcoholism in the overall sample: the authors reported that over 50% of the sample was chemically dependent There were, however, no significant MMPI differences between the dependent and nondependent sexual offenders.

In terms of types of child molester, Erickson et al. reported there were no significant differences on two-point codes between molesters of female and male children. The Pd-D/D-Pd and Pd-Sc/Sc-Pd profiles were significantly more common among extrafamilial versus intrafamilial child molesters. Finally, the Pd-D/D-Pd and Pd-Sc/Sc-Pd codes were the two most common profiles, both occurring at 15.3% frequency, among sexual offense recidivists (those who had previously been convicted of a sexual offense). Recidivists also had more elevated MMPI scales in general. Although these two point code findings in a large heterogeneous sample of sex offenders are impressive, the authors are correct in emphasizing that even the most common two-part codes were still of relatively low frequencies.

Quinsey, Arnold and Pruesse (1980) found that rapists and extrafamilial child molesters tended to have elevations on scales Pd, D, and Sc. They did not, however, find that the child molesters and rapists were significantly different from each other. The sexual offenders were also not different from the other groups of offenders at the maximum security psychiatric hospital where they selected their comparison samples (intrafamial murderers, extrafamilial murderers, arsonists, and property offenders.)

The most methodologically sound studies in this area (aside from that of Erickson et al., 1987) are likely those conducted by Hall and his colleagues (Hall, 1988; Hall, 1989b; Hall, Maiuro, Vitaliano & Proctor, 1986; Hall & Proctor, 1987).

In a sample of 406 male pedophiles, the mean Pd and Sc scales were both significantly elevated, but no single two-point code was predominant. In fact, all of the 10 MMPI clinical scales exceeded a t-score of 60 in the mean profile of the sample. The mean Pd, Sc and D scales were all at t = 70 or above, but not one individual in the sample had these three scales elevated exclusively. The two-point profiles were scattered. The most common codes were Pd-Sc/Sc-Pd and Psychasthenia-Sc/Sc-Psychasthenia, but these codes were only found in 7.1% and 6.4% of the sample, respectively. There was a small but significant difference between heterosexual and homosexual pedophiles on the Mf scale. They indicate that the most common findings is that of emotional distress, as evidenced by the fact that two-thirds of the sample had three or more scale elevations. Most importantly, the results strongly suggest that pedophiles, like sex offenders in general, are a heterogeneous population with regards to MMPI profiles. This was again confirmed in a study of 80 pedophiles by Hall (1989b).

Hall (1988) examined the recidivism of 342 of the sex offenders studied by Hall and Proctor (1987). In general, actuarial predictions using the MMPI outperformed clinical judgement predictions for recidivism. The Mf scale was part of a significant actuarial equation that predicted sexual re-offenses against adults. The Hypochondriasis, D and Pd scales were part of a significant actuarial prediction of nonsexual violent offenses. The Lie and D scales were part of a significant actuarial equation to predict nonsexual nonviolent offenses. Finally, and perhaps most importantly, none of the MMPI scales (nor clinical judgement) significantly predicted sexual re-offending against children.

Hanson et al. (1990a) found that child molesters in treatment showed significant improvements on all the MMPI scales except Ma (and L). However, only the posttest Mf scores predicted recidivism for a sexual or violent offense ($\underline{r} = .20$, $\underline{p} < .05$). The interpretation of this finding is difficult since the effect was weak and many comparisons were conducted. Nevertheless, since the Mf scale has predicted recidivism in other studies, the usefulness of the Mf scale with sexual offenders deserves further consideration.

Conclusions and Recommendations

The results from MMPI profile studies indicate that sex offenders are a heterogeneous population. Many of the significant differences that have been observed between sex offenders and normals and between pedophiles and normals are likely due to extraneous factors such as emotional distress due to incarceration or stigma. Almost all of the MMPI studies have found that sex offenders, including pedophiles, have elevated Pd scales. Yet differences among types of sex offenders are scattered and often not replicated. Sometimes homosexual and heterosexual offenders significantly differ on the Mf scale. Sometimes pedophiles have elevated Si and Sc scales and lower Hypomania scores compared to other types of offenders. However, these differences are not large and are not consistent across studies. What is perhaps more important than demonstrating overall group differences is the discriminatory power of MMPI scales on a case-by-case basis. Few studies have employed this type of analysis and the results are not encouraging.

In terms of the clinical significance of the MMPI based on the research to date, it is clear that few pedophiles can be reliably differentiated at assessment from other types of offenders and this is doubly true for predicting paedophile recidivism. More research is still needed on the sexual deviantions scales of the MMPI (Pedophila & Sexual Deviation) rather than the clinical scales (with the possible exception of MF). Clinicians may want to use the MMPI as a general measure of psychopathology and personality, but it provides little information that is relevant to the identification of sexual offender or that is useful in predicting recidivism. It remains to be seen whether the new content scales of the MMPI-2, such as the Antisocial Practices and Negative Treatment Indicators scales, will prove useful with sexual offenders (Graham, 1990).

Conclusions and future directions

Most of the measures reviewed appear to have limited utility in the risk management of sexual offenders. Sexual offenders are difficult to study. Many different factors contribute to sexual offending. As well, sexual offenders are not motivated to self-disclose their motivation for offending, or even admit to their offenses. Nevertheless, the importance of accurately assessing sexual offenders justifies further work in this complex area. Although there are some measures that appear preferable to others, considerably more research needs to be directed towards developing better measures.

The specific measures that clinicians and researchers will want to adopt depend, in part, on the goals of their treatment or research projects. Nevertheless, some general comments about the merits of various measures are in order. These recommendations should be considered tentative since the research is not developed enough to justify firm conclusions.

Of the various sexual history measures, the Clarke SHQ (Langevin et al., 1990) appears to be the most useful with sexual offenders. It covers a broad range of relevant topics and is reasonably easy to administer. It need not be formally scored, but is perhaps most useful as a checklist of sexual acts that the subjects are willing to admit. Since it is a history measure, it is not useful as a treatment outcome measure - except in programs that measure progress by increased willingness to self-disclose deviant sexual activities.

The Abel/Becker/Laws card sort procedure appears to be the most promising self-report measure of deviant sexual interests. Although the research is limited, its reported ability to elicit deviant interests from offenders who deny such interests in interviews is valuable. It can also be used as a treatment outcome measure in programs that aim to modify deviant sexual preferences.

Abel et al.'s (1989) Cognitions measure appears to be a useful attitude measure to use with child molesters. The six items from the "Child-adult sex helps the child" factor are easy to administer, internally consistent, and there is even some evidence that they differentiate child molesters from nonoffenders. These items are from the first rotated factor, however. An alternate approach to simplifying the scale would be to select items from the first unrotated factor, or, equivalently, the items with the highest item-total correlations.

None of the attitude measures for rapists have received adequate research support to be recommended. The basic theory that rapists will have attitudes tolerant of rape appears credible, but none of the available measures have been adequately validated in applied settings. Most of the commonly used scales tend to lack internal consistency and be unable to differentiate identified rapist from nonsexual offenders or low SES community controls. Although it has not been used with sexual offenders, one of the more promising of these scales is Mosher's Callous Sex Attitudes from his Hypermasculinity scale. The items on this scale appear to match the criminal sentiments considered to be associated with criminal recidivism generally (Andrews & Wormith, 1984; Wormith, 1984).

Many of the sex role scales appear to be reliable and valid. These scales, however, have not been able to differentiate identified sexual offenders from nonoffenders in applied settings (see Overholser & Beck, 1986, for an exception), although they are able to predict self-reported sexual aggression in college students (e.g., Malamuth, 1986). The problem with the sex role measures may be that offenders are unwilling to disclose their sexist attitudes; however, the theory linking sexist attitudes to rape may also be weak. Conventional sex roles may condone male domination, but contrary to some feminist positions (e.g., Burt, 1980), there need not be a direct progression from traditional sex roles to rape. It is easy to imagine highly conservative men who abhor sexual assault; their attitudes towards women may be demeaning in a variety of ways, but they need not condone rape. Rather than traditional sex roles per se, it may be power-oriented, authoritarian sexual values that are associated with rape, although authoritarianism and traditional sex role values would be expected to be related. Consequently, there may be little direct relationship between rape and many of the sex role issues addressed in the available questionnaires (e.g., whether swearing is more offensive in a man than a women, or whether women can make good engineers). The measures that are most likely to differentiate sexual offenders from nonsexual offenders are those that address sexist attitudes specific to heterosexual relationships, e.g., male dominance as justified in obtaining sex from women. Hall et al's (1986) Heterosexual Relationships Scale is one of the few scales designed to assess sexist attitudes specific to sexual interactions, but the psychometric adequacy of their measure remains questionable. Researchers who are interested in omnibus measures of conservative sex roles will likely find the more subtle recent measures, such as the BIAS (Jean & Reynolds, 1980) or the Index of Sex-Role Orientation (Dreyer, et al., 1981), preferable to the now somewhat dated Attitudes Towards Women scale (Spence et al., 1973).

The most established of the sexual dysfunction measures is the Derogatis Sexual Functioning Inventory (Derogatis, 1980). It covers important content areas and appears to have acceptable levels of reliability and validity. Although identified sexual offenders report dissatisfaction with their sexual life, it is unclear how sexual dysfunction is related to sexual offending. It is likely to be most meaningful with those offenders lacking deviant sexual interests (e.g., most incest offenders - Lang & Frenzel, 1989). The DSFI is not designed to assess men with sexual anomalies.

There are several areas that are not well covered by existing measures. Masturbation activity is important in the assessment of sexual offenders, but there are no measures appropriate for use with sexual offenders. The closest measure is the Wilson (1978) Fantasy measure, but it addresses only masturbation fantasies. A masturbation measure for sexual offenders should address frequency and satisfaction as well.

Another area that is not well assessed is that of childhood sexual trauma. The role of sexual abuse in the etiology of sexual offending is debatable (Garland & Dougher, 1990; Hanson & Slater, 1988); nevertheless, sexually abused child molesters often have problems that are not found in nonabused offenders (Hanson, 1991; Langevin et al., 1989). A related area that is not assessed is that of deviant models. The role of deviant models in criminal behaviour is well established (e.g., Andrews, 1980; Sutherland, 1947). It is quite plausible, consequently, that the chances of committing sexual offenses would also increase through association with peers who condone such acts.

Assessing abilities, rather than attitudes or history, is another assessment approach worthy of research attention. The advantage of skill measures is that they are not amenable to self-presentation biases. The difficulty with skill measures is identifying the specific deficits that need to be assessed. Several recently developed skill measures for sexual offenders appear promising, although none of these measures have received sufficient research to adequately evaluate them.

One skill approach is the situational competency test developed by Miner et al. (1989). This measure involves describing situations in which there is a high risk for reoffending, and then asking the offender to provide his best coping response. For example, a child molester may be asked what he would do if he was at a children's birthday party and the only other adult in attendance stated that she was going to the store to get ice cream. The adequacy of the coping responses can be evaluated on several dimensions, including response time latency, the feasibility of the plan, and its complexity. Preliminary results from the Atascadero program indicate that it is able to detect increases in coping skills among the treated sexual offenders (Miner et al., 1990). Although Miner et al.'s test is presented on audiotape, it would also be possible to create a written version.

Situational competency tests could be used in conjunction with self-assessments of the types of situations that specific offenders consider to be high risk. One possible situational risk assessment measure is the one developed by Stermac, Glancy and Davidson (1989). Stermac et al.'s scale is closely modelled after Annis' (1982) Inventory of Drinking Situations.

Another promising approach is the cognition/correction procedure developed by N. Couture at the Centre hospitalier Robert-Giffard in Quebec City (Isabelle Pronovost, personal communication). The offender is first asked to provide accounts of his offenses. The cognitive distortions in his account are identified by the therapist (e.g., she wanted it) and the offender is asked to provide corrections to such distortions (e.g., that is only my interpretation). Next, the offender is given a list of cognitive distortions provided by other offenders and asked to provide corrections to them. The speed and adequacy of the offenders' corrections can then be evaluated. The cognition/correction method has not been evaluated, but it is a promising approach that may be less susceptible to self-presentation bias than other cognitions scales, such as Abel et al.'s (1989).

Another strategy for assessing offenders' deficits follows from McFall's (1990) social skills models. One reason that some offenders may sexually assault their victims is that they misinterpret the social cues and fail to perceive their actions as abuse. For example, rapists may think that their victims really wanted and enjoyed the sexual activity, or child molesters may misinterpret children's affection as sexual interest. Tests of this type would ask offenders for correct interpretations of complex and/or competing social cues. Preliminary data with one such test for child molesters (Hanson, 1991) provided initial support for this approach. When asked to rate the abusiveness of a set of 15 adult-child interactions, the overall ratings of abusiveness were not different for child molesters, sexual abuse experts, or community controls; however, the accuracy of the child molesters' ratings were significantly worse than the ratings of the control subjects.

Self-monitoring measures for sexual offenders are also worth developing. These measure could involve standard formats for daily recording of sexual activities or fantasies. For example, an offender could be asked to record basic information each time he was sexually stimulated. This information could include the time, place, and people present, accompanying thoughts and feelings, and any actions taken to subvert deviant sexual impulses. Such self-monitoring strategies have already been successfully used in the assessment and treatment of anxiety, stress and depression (e.g., Barlow, 1988; Beck, Rush, Shaw & Emery, 1979; Meichenbaum, 1977). The utility of self-monitoring methods with sexual offenders has not been evaluated. Nevertheless, they could be valuable for offenders at advanced stages of treatment who are motivated to prevent further offenses.

The general conclusion of this review is that further work is necessary to improve upon the existing self-report sexuality measures for sexual offenders. Some useful scales are available, but many areas are not well assessed by existing measures. Clinicians and researchers are often better off constructing new measures that are suited to their purposes than relying on the measures that are available. Treatment programs, in particular, would benefit from developing their own treatment-specific skill measures that can be used to evaluate improvement in their form of treatment. Continued research will, hopefully, identify the general approaches and specific measures that are most useful in the risk assessment of sexual offenders.

References

Abel, G. G., Becker, J. V., & Cunningham-Rathner, J. (1984). Complications, consent, and cognitions in sex between children and adults. International Journal of Law and Psychiatry, 7, 89-103.

Abel, G. G., Becker, J. V., Cunningham-Rathner, J., Rouleau, J.L., Kaplan, M., & Reich, J. (1981). The treatment of child molesters. Unpublished manuscript. Sexual Behavior Clinic, Columbia University, New York.

Abel, G. G., Becker, J. V., Mittelman, M., Cunningham-Rathner, J., Rouleau, J. L., & Murphy, W. D. (1987). Self-reported sex crimes of nonincarcerated paraphiliacs. Journal of Interpersonal Violence, 2, 3-25.

Abel, G. G., Cunningham-Rathner, J., Becker, J. V., & McHugh. (1983, December). Motivating sex offenders for treatment with feedback of their psychophysiologic assessment. Presentation at the World Congress of Behavior Therapy, Washington, D.C.

Abel, G. G., Gore, D. K., Holland, C. L., Camp, N., Becker, J.V., & Rathner, J. (1989). The measurement of cognitive distortions of child molesters. Annals of Sex Research, 2, 135-152.

Abramson, P. R., & Mosher, D. L. (1975). Development of a measure of negative attitudes towards masturbation. Journal of Consulting and Clinical Psychology, 43, 485-490.

Alford, J. M., & Brown, G. E. (1985). Virgins, whores, and bitches: Attitudes of rapists toward women and sex. Corrective and Social Psychiatry, 31, 58-61.

Allen, R. M., & Haupt, T. D. (1966). The Sex Inventory: Test-retest reliabilities of scale scores and items. Journal of Clinical Psychology, 22, 375-378.

Anastasi, A. (1982). Psychological testing. New York: MacMillan.

Andrews, D. A. (1980). Some experimental investigations of the principles of differential association through deliberate manipulation of the structure of service systems. American Sociological Review, 45, 448-462.

Andrews, D. A., & Wormith, J. S. (1984). Criminal sentiments and criminal behavior. Programs Branch User Report. Ottawa: Solicitor General of Canada.

Annis, H. M. (1982). Inventory of Drinking Situations. Available from Addiction Research Foundation, 33 Russell Street, Toronto, Ontario, M5S 2S1.

Archer, J., & Rhodes, C. (1989). The relationship between gender-related traits and attitudes. British Journal of Social Psychology, 28, 149-157.

Armentrout, J. A., & Hauer, A. L. (1978). MMPIs of rapists of adults, rapists of children, and non-rapist sex offenders. Journal of Clinical Psychology, 34, 330-332.

Ashton, N. L. (1982). Validation of Rape Myth Acceptance Scale. Psychological Reports, 50, 252.

Barlow, D. H. (1988). Anxiety and its disorders. New York: Basic Books.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, S. (1979). Cognitive therapy of depression: A treatment manual. New York: Guilford.

Bem, S. L. (1981). Bem Sex Role Inventory: Professional manual. Palo Alto, CA: Consulting Psychologist Press.

Bentler, P. M. (1968). Heterosexual behavior assesment - I. Males. Behavior Research and Therapy, 6, 21-25.

Briere, J., Malamuth, N., & Check, J. V. P. (1985). Sexuality and rape-supportive beliefs. International Journal of Women's Studies, 8, 398-403.

Briere, J., & Runtz, M. (1989). University males' sexual interest in children: Predicting potential indices of "pedophilia" in a nonforensic sample. Child Abuse and Neglect, 13, 65-75.

Burt, M. R. (1980). Cultural myths and supports for rape. Journal of Personality and Social Psychology, 38, 217-230.

Burt, M. R. (1983). Justifying personal violence: A comparison of rapists and the general public. Victimology, 8, 131-150.

Burt, M. R., Albin, R. S. (1981). Rape myths, rape definitions, and probability of conviction. Journal of Applied Social Psychology, 11, 212-230.

Buss, A. H., & Durkee, A. (1957). An inventory for assessing different kinds of hostility. Journal of Consulting Psychology, 21, 343-349.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Canada. (1990). *The management and treatment of sex offenders*. Ottawa: Minister of Supply and Services.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Chaudhuri, A., & Mukerjee, R. (1988). *Randomized response: Theory and techniques*. New York: Marcel Dekker.

Check, J. V. P. (1985). The Hostility Towards Women Scale (Doctoral dissertation, University of Manitoba, 1984). *Dissertation Abstracts International, 45(12)*.

Check, J. V. P., & Malamuth, N. M. (1983). Sex role stereotyping and reactions to depictions of stranger versus acquaintance rape. *Journal of Personality and Social Psychology, 45*, 344-356.

Cliff, N. (1988). The eigenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin, 103*, 276-279.

Colligan, R. C., Osborne, D., Swenson, W. M., & Offord, K. P. (1984). The MMPI: Development of contemporary norms. *Journal of Clinical Psychology, 40*, 100-107.

Cortoni, F., Gordon, A., Malcolm, B., & Ellerby, L. (1991). The Attitudes Towards Sex With Children Scale: Preliminary results. *Canadian Psychology/Psychologie canadienne, 23 (2a)*, 229.

Cotten-Hustan, A. L. (1983). Comparison of sex offenders and non-offenders on attitudes towards masturbation and female fantasy as related to participation in human sexuality sessions. *Journal of Offender Counselling, Services, & Rehabilitation, 8*, 13-26.

Cowden, J. E., & Pacht, A. R. (1969). The Sex Inventory as a classication instrument for sex offenders. *Journal of Clinical Psychology, 23*, 53-57.

Craig, M. E., Kalichman, S. C., & Follingstad, D. R. (1989). Verbal coercive sexual behavior among college students. *Archives of Sexual Behavior, 18*, 421-434.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 24, 349-354.

Davis, C. M., Yarber, W. L., & Davis, S. L. (1988). Sexuality related measures: A compendium. Lake Mills, IA: Graphic Publishing Co.

Deitz, S. R., Littman, M., & Bentley, B. J. (1984). Attribution of responsibility for rape: The influence of observer empathy, victim resistance, and victim attractiveness. Sex Roles, 10, 261-280.

Deitz, S. R., Tiemann-Blackwell, K., Daley, P. C., & Bentley, B. J. (1982). Measurement of empathy towards rape victims and rapists. Journal of Personality and Social Psychology, 43, 372-384.

Demare, D., Briere, J., & Lips, H. M. (1988). Violent pornography and self-reported likelihood of sexual aggression. Journal of Research in Personality, 22, 140-153.

Derogatis, L. R. (1980). Psychological assessment of psychosexual functioning. Psychiatric Clinics of North America, 3, 113-131.

Derogatis, L. R., & Melisaratos, N. (1979). The DSFI: A multidimensional measure of sexual functioning. Journal of Sex and Marital Therapy, 5, 244-281.

Dreyer, N. A., Woods, N. F., & James, S. A. (1981). ISRO: A scale to measure sex-role orientation. Sex Roles, 7, 173-182.

Dwyer, S. M., & Amberson, J. I. (1989). Behavioral patterns and personality characteristics of 56 sex offenders: A preliminary study. Journal of Psychology and Human Sexuality, 2, 105-118.

Erickson, W. D., Luxenberg, M. G., Walbek, N. H., & Seely, R. K. (1987). Frequency of MMPI two-point code types among sex offenders. Journal of Consulting and Clinical Psychology, 55, 566-570.

Eysenck, H. J. (1971). Personality and sexual adjustment. British Journal of Psychiatry, 118, 593-608.

Eysenck, H. J. (1978). Sex and personality. London: Open Books.

Feild, H. S. (1978). Attitudes toward rape: A comparitive analysis of police, rapists, crisis counselors and citizens. Journal of Personality and Social Psychology, 36, 156-179.

Finkelhor, D. (1984). Child sexual abuse: New theory and research. New York: Free Press.

Finkelhor, D., & Lewis, I. A. (1988). An epidemiological approach to the study of child molesting. In R. A. Prentky and V. L. Quinsey (Eds.), Human sexual aggression: Current perspectives (pp. 64-78). New York: New York Academy of Sciences.

Fisher, G., & Howell, L. M. (1970). Psychological needs of homosexual pedophiles. Diseases of the Nervous System, 31, 623-625.

Fox, J. A., & Tracy, P. E. (1986). Randomized response: A method for sensitive surveys. Beverly Hills: Sage.

Freund, K., Langevin, R., Laws, R., & Serber, M. (1974). Femininity and preferred partner age in homosexual and heterosexual males. British Journal of Psychiatry, 125, 442-446.

Garland, R. J., & Dougher, M. J. (1990). The abused/abuser hypothesis in child sexual abuse: A critical review of theory and research. In J. R. Feierman (Ed.), Adult human sexual behavior with children and adolescents: Biosocial dimensions (pp. 488-509). New York: Aldine de Grayter.

Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). Measurement theory for the behavioral sciences. San Francisco: W. H. Freeman.

Gillis, J. R. (1991). The rapist as a sexually deviant offender. Unpublished doctoral dissertation, Queens University, Kingston.

Gore, D. K. (1988). Measuring the cognitive distortions of child molesters: Psychometric properties of the cognition scale (Doctoral dissertation, Georgia State University, 1988).

Gorsuch, R. L. (1983). Factor analysis (2nd. ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gosselin, C., & Wilson, G. (1980). Sexual variations: Fetishism, sadomasochism, and transvestism. New York: Simon & Schuster.

Graham, J. R. (1990). MMPI-2: Assessing personality and psychopathology. Oxford: Oxford University Press.

Greendlinger, V. A. (1985). Dispositional and situational variables as predictors of rape proclivity in college males. (Doctoral dissertation, State University of New York at Albany, 1985).

Greendlinger, V. A., & Byrne, D. (1987). Coercive sexual fantasies of college men as predictors of self-reported likelihood to reape and overt sexual aggression. Journal of Sex Research, 23, 1-11.

Hall, E. R., Howard, J. A., & Boezio, S. L. (1986). Tolerance of rape: A sexist or antisocial attitude? Psychology of Women Quarterly, 10, 101-118.

Hall, G. C. N. (1988). Criminal behavior as a function as clinical and actuarial variables in a sexual offender population. Journal of Consulting and Clinical Psychology, 56, 773-775.

Hall, G. C. N. (1989a). Self-reported hostility as a function of offense characteristic and response style in a sexual offender population. Journal of Consulting and Clinical Psychology, 57, 306-308.

Hall, G. C. N. (1989b). WAIS-R and MMPI profiles of men who have sexually assaulted children: Evidence of limited utility. Journal of Personality Assessment, 53, 404-412.

Hall, G. C. N. (1990). Prediction of sexual aggression. Clinical Psychology Review, 10, 229-245.

Hall, G. C. N., & Proctor, W. C. (1987). Criminological predictors of recidivism in a sexual offender population. Journal of Consulting and Clinical Psychology, 55, 111-112.

Hall, G. C. N., Maiuro, R. D., Vitaliano, P. P., & Proctor, W. C. (1986). The utility of the MMPI with men who have sexually assaulted children. Journal of Consulting and Clinical Psychology, 54, 493-496.

Hamilton, M., & Yee, J. (1990). Rape knowledge and propensity to rape. Journal of Research in Personality, 24, 111-122.

Hanson, R. K. (1991). Characteristics of sex offenders who were sexually abused as children. In R. Langevin (Ed.), Sex offenders and their victims: New research findings (pp. 77-85). Toronto: Juniper Press.

Hanson, R. K. (1991). The Sexual Abuse Empathy Test: Evaluation of Pilot Testing. Report for contract # 1512-HA01300-9001/824 from the Ministry Secretariat, Solicitor General Canada.

Hanson, R. K., & Slater, S. (1988). Sexual victimization in the history of child molesters: A review. Annals of Sex Research, 1, 485-499.

Hanson, R. K., Steffy, R. A., & Gauthier, R. (1990a, October). Long-term follow-up of treated child molesters in Canada. Presentation at the Ninth Annual Clinical and Research Conference on the Assessment and Treatment of Sexual Abusers, Their Families and Victims, Toronto.

Hanson, R. K., Steffy, R. A., & Gauthier, R. (1990b, August). Recidivism of child molesters 20 years after treatment. Presentation at the American Psychological Association Convention, Boston.

Hare, R. D., Harpur, T. J., Hakstian, A. R., Forth, A. E., Hart, S. D., & Newman, J. P. (1990). The Revised Psychopathy Checklist: Reliability and factor structure. Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2, 338-341.

Haupt, T. D., & Allen, R. M. (1966). A multivariate analysis of variance of scale scores on the Sex Inventory Male Form. Journal of Clinical Psychology, 22, 387-395.

Hodkinson, S. M. (1990). The perception of consent to sexual contact in children: Child molesters versus non-offenders. Unpublished masters thesis, York University, Toronto.

Horley, J. (1988). Cognitions of child sexual abusers. Journal of Sex Research, 25, 542-545.

Howard, J. A. (1984). Societal influences on attribution: Blaming some victims more than others. Journal of Personality and Social Psychology, 47, 494-505.

Howells, K. (1979). Some meanings of children for pedophiles. In M. Cook & G. Wilson (Eds.), Love and attraction: An international conference (pp. 519-526). Oxford: Pergamon.

Howells, K., & Wright, E. (1978). The sexual attitudes of aggressive sexual offenders. British Journal of Criminology, 18, 170-174.

Intons-Peterson, M. J., Roskos-Ewoldsen, B., Thomas, L., Shirley, M., & Blut, D. (1989). Will educational materials reduce negative effects of exposure to sexual violence? Journal of Social and Clinical Psychology, 8, 256-275.

Jackson, D. N. (1974). Personality Research Form manual - Form E. Goshen, NY: Research Psychologist Press.

Jean, P. J., & Reynolds, C. R. (1980). Development of the Bias in Attitudes Survey: A sex-role questionnaire. The Journal of Psychology, 104, 269-277.

Jean, P. J., & Reynolds, C. R. (1984). Sex and attitude distortion: Ability of females and males to fake liberal and conservative positions regarding changing sex roles. Sex Roles, 10, 805-815.

Joreskog, K. G., & Sorbom, D. (1986). LISREL-6. Mooreville, IN: Scientific Software.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20, 141-151.

Kalichman, S. C. (1991). Psychopathology and personality characteristics of criminal sexual offenders as a function of victim age. Archives of Sexual Behavior, 20, 187- 197.

Kalichman, S. C., Craig, M. E., Shealy, L., Taylor, J. Szymanowski, D., & McKee, G. (1989). An empirically derived typology of adult rapists based on the MMPI: A cross-validation study. Journal of Psychology and Human Sexuality, 2, 165-182.

Kalichman, S., Szymankowski, D., McKee, G., Taylor, J., & Craig, M. E. (1989). Cluster analytically derived MMPI profile subgroups in incarcerated adult rapists. Journal of Clinical Psychology, 45, 149-155.

Kelly, G. A. (1955). The psychology of personal constructs (Vols. 1 & 2). New York: Norton.

Koss, M. P., & Gidycz, C. A. (1985). Sexual Experiences Survey: Reliability and validity. Journal of Consulting and Clinical Psychology, 53, 422-423.

Koss, M. P., Leonard, K. E., Beezley, D. A., & Oros, C. J. (1985). Nonstranger sexual aggression: A discriminant analysis of the psychological characteristics of undetected offenders. Sex Roles, 12, 981-991.

Koss, M. P., & Oros, C. J. (1982). Sexual Experiences Survey: A research instrument investigating sexual aggression and victimization. Journal of Consulting and Clinical Psychology, 50, 455-457.

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. Psychometrika, 2, 151-160.

Lang, R. A., & Frenzel, R. R. (1989). Identifying sexual preferences among intrafamilial and extrafamilial child sexual abusers. Annals of Sex Research, 2, 255-276.

Lang, R. A., Langevin, R., Checkley, K. L., & Pugh, G. (1987). Genital exhibitionism: Courtship disorder or narcissism? Canadian Journal of Behavioural Science, 19, 216-232.

Langevin, R., Paitich, D., Freeman, R., Mann, K., & Handy, L. (1978). Personality characteristics and sexual anomalies in males. Canadian Journal of Behavioural Science, 10, 222-238.

Langevin, R., Paitich, D., Russon, A., Handy, L., & Langevin, A. (1990). Clarke Sex History Questionnaire for Males manual. Oakville, Ontario: Juniper Press.

Langevin, R., Wright, P., & Handy, L. (1989). Characteristics of sex offenders who were sexually victimized as children. Annals of Sex Research, 2, 227-253.

Langevin, R., Wright, P., & Handy, L. (1990a). Use of the MMPI and its derived scales with sex offenders: I. Reliability and validity studies. Annals of Sex Research, 3, 245-291.

Langevin, R., Wright, P., & Handy, L. (1990b). Use of the MMPI and its derived scales with sex offenders: II. Reliability and criterion validity. Annals of Sex Research, 3, 453-486.

Laws, D. R. (1984). The assessment of diverse sexual behavior. In K. Howells (Ed.), The psychology of sexual diversity (pp. 111-134). Oxford: Basil Blackwell.

Laws, D. R. (Ed.). (1989). Relapse prevention with sex offenders. New York: Guilford Press.

Laws, D. R., Greenbaum, P. E., Osborn, C. A., & Murrin, M. R. (1989). Classification of child molesters by physiological assessment of sexual arousal and a self-report measure of sexual preference. Unpublished data, Florida Mental Health Institute, Tampa, FL.

Linz, D. G., Donnerstein, E., & Penrod, S. (1988). Effects of long-term exposure to violent and sexually degrading depictions of women. Journal of Personality and Social Psychology, 55, 758-768.

Lipton, D. N., McDonel, E. C., & McFall, R. M. (1987). Heterosocial perception in rapists. Journal of Consulting and Clinical Psychology, 55, 17-21.

Lisak, D., & Roth, S. (1988). Motivational factors in nonincarerated sexually aggressive men. Journal of Personality and Social Psychology, 55, 795-802.

Lisak, D., & Roth, S. (1990). Motives and psychodynamicx of self-reported, unincarcerated rapists. American Journal of Orthopsychiatry, 60, 268-280.

Lottes, I. L. (1988). Rape Supportive Attitude Scale. In C. M. Davis, W. L. Yarber, S. L., & Davis (Eds.) Sexuality related measures: A compendium. Lake Mills, IA: Graphic Publishing Co.

Malamuth, N. (1986). Predictors of naturalistic sexual aggression. Journal of Personality and Social Psychology, 50, 953-962.

Malamuth, N. (1989a). The Attraction to Sexual Aggression Scale: Part One. The Journal of Sex Research, 26, 26-49.

Malamuth, N. (1989b). The Attraction to Sexual Aggression Scale: Part Two. The Journal of Sex Research, 26, 324-354.

Malamuth, N., & Check, J. V. P. (1983). Sexual arousal to rape depictions: Individual differences. Journal of Abnormal Psychology, 92, 55-67.

Malamuth, N., Check, J. V. P., & Briere, J. (1986). Sexual arousal in response to aggression: Ideological, aggressive and sexual correlates. Journal of Personality and Social Psychology, 50, 330-340.

March, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. Psychological Bulletin, 103, 391-410.

Marsh, J. T., Hilliard, J., & Liechti, R. (1955). A sexual deviation scale for the MMPI. Journal of Consulting Psychology, 19, 55-60.

Marolla, J., & Scully, D. (1986). Attitudes toward women, violence, and rape: A comparison of convicted rapists and other felons. Deviant Behavior, 7, 337-355.

Marques, J. K., Day, D. M., Nelson, C., & Miner, M. H. (1989). The Sex Offender Treatment and Evaluation Project: California's relapse prevention program. In D. R. Laws, (Ed.), Relapse prevention with sex offenders (pp. 247-267). New York: Guilford Press. see page 263.

Marshall, W. L., Laws, D. R., & Barbaree, H. E. (Eds.). (1990). Handbook of sexual assault: Issues, theories and treatment of the offender. New York: Plenum.

McFall, R. M. (1990). The enhancement of social skills: An information-processing analysis. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), Handbook of sexual assault: Issues, theories and treatment of the offender (pp. 311-330). New York: Plenum.

Meichenbaum, D. (1977). Cognitive-behavior modification: An integrative approach. New York: Plenum Press.

Miner, M. H., Day, D. M., & Nafpaktitis, M. K. (1989). Assessment of coping skills: Development of a situational competency test. In R. D. Laws (Ed.), Relapse prevention with sex offenders (pp. 127-136). New York: Guilford Press.

Miner, M. H., Marques, J. K., Day, D. M., & Nelson, C. (1990). Impact of relapse prevention in treating sex offenders: Preliminary findings. Annals of Sex Research, 3, 165-185.

Mosher, D. L. (1966). The development and multitrait-multimethod matrix analysis of three measures of guilt. Journal of Consulting Psychology, 30, 25-29.

Mosher, D. L. (1971a). Psychological reactions to pornographic films. In Technical report on the commission on obscenity and pornography: Vol. 8, Erotica and social behavior. Washington, DC: U.S. Govt. Printing Office.

Mosher, D. L. (1971b). Sex callousness toward women. In Technical report on the commission on obscenity and pornography: Vol. 8. Erotica and social behavior. Washington, DC: U.S. Govt. Printing Office.

Mosher, D. L. (1988a). Aggressive Sexual Behavior Inventory. In C. M. Davis, W. L. Yarber, S. L., & Davis (Eds.) Sexuality related measures: A compendium (pp. 9-10). Lake Mills, IA: Graphic Publishing Co.

Mosher, D. L. (1988b). Negative Attitudes Towards Masturbation. In C. M. Davis, W. L. Yarber, S. L., & Davis (Eds.) Sexuality related measures: A compendium (pp. 227-229). Lake Mills, IA: Graphic Publishing Co.

Mosher, D. L., & Anderson, R. D. (1986). Macho personality, sexual aggression, and reactions to guided imagery of realistic rap. Journal of Research in Personality, 20, 77-94.

Mosher, D. L., & Sirkin, M. (1984). Measuring a macho personality constellation. Journal of Research in Personality, 18, 150-163.

Muehlenhard, C. L., Friedman, D. E., & Thomas, C. M. (1985). Is date rape justifiable? The effects of dating activity, who initiated, who paid, and men's attitudes toward women. Psychology of Women Quarterly, 9, 297-310.

Muliak, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. Psychological Bulletin, 105, 430-445.

Murphy, W. D. (1990). Assessment and modification of cognitive distortions in sex offenders. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), Handbook of sexual assault: Issues, theories and treatment of the offender (pp. 331-342). New York: Plenum.

Nadler, E., & Morrow, W. (1959). Authoritarian attitudes towards women, and their correlates. Journal of Social Psychology, 49, 113-123.

Nichols, H. R., & Molinder, I. (1984). Multiphasic Sex Inventory. (Available from Nichols & Molinder, 437 Bowes Drive, Tacoma, WA 98466).

Overholser, J. C., & Beck, S. (1986). Multimethod assessment of rapists, child molesters, and three control groups on behavioral and psychological measures. Journal of Consulting and Clinical Psychology, 54, 682-687.

Paitich, D., Langevin, R., Freeman, R., Mann, K., & Handy, L. (1977). The Clarke SHQ: A clinical sex history questionnaire for males. Archives of Sexual Behavior, 6, 421-435.

Panton, J. H. (1979). MMPI profile configurations associated with incestuous and non-incestuous child molesting. Psychological Reports, 45, 335-338.

Parry, G. (1983). A British version of the Attitudes Towards Women Scale (AWS-B). British Journal of Social Psychology, 22, 261-263.

Pawlak, A. E., Boulet, J. R., & Bradford, J. M. W. (1991). Discriminant analysis of a sexual-functioning inventory with intrafamilial and extrafamilial child molesters. Archives of Sexual Behavior, 20, 27-34.

Peterson, D. L., & Pfost, K. S. (1989). Influence of rock videos on attitudes of violence against women. Psychological Reports, 64, 319-322.

Pryor, J. B. (1987). Sexual harassment proclivities in men. *Sex Roles, 17*, 269-290.

Quackenbush, R. L. (1989). A comparison of androgynous masculine sex-typed, and undifferentiated males on dimensions of attitudes toward rape. *Journal of Research in Personality, 23*, 318-342.

Quinsey, V. L. (1984). Sexual aggression: Studies of offenders against women. In D. N. Weisstub (Ed.), *Law and mental health: International perspectives: Vol. 1* (pp. 84-123). New York: Pergamon.

Quinsey, V. L. (1986). Men who have sex with children. In D. N. Weisstub (Ed.), *Law and mental health: International perspectives: Vol. 2* (p[illegible]40-172). New York: Pergamon.

Quinsey, V. L., Arnold, L. S., & Pruesse, M. G. (1980). MMPI profiles of men referred for a pretrial psychiatric assessment as a function of offense type. *Journal of Clinical Psychology, 36*, 410-417.

Quinsey, V. L., & Earls, C. M. (1990). The modification of sexual preferences. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories and treatment of the offender* (pp. 279-295). New York: Plenum.

Rombough, S., & Ventimiglia, J. C. (1981). Sexism: A tri-dimensional phenomenon. *Sex Roles, 7*, 747-755.

Salter, A. C. (1988). *Treating child sex offenders and victims*. Newbury Park, CA: Sage.

Sattem, L., Savells, J., & Murray, E. (1984). Sex-role stereotypes and commitment of rape. *Sex Roles, 11*, 849-860.

Saunders, E. J. (1988). A comparative study of attitudes towards child sexual abuse among social work and judicial system professionals. *Child abuse and neglect, 12*, 83-90.

Schiavi, R. C., Derogoatis, L. R., Kuriansky, J., O'Connor, D., & Sharpe, L. (1979). The assessment of sexual function and marital interaction. *Journal of Sex and Marital Therapy, 5*, 169-224.

Schmitt, N., & Stults, D. M. (1986). Methodology review: Analysis of multitrait-multimethod matrices. *Applied Psychological Measurement, 10*, 1-22.

Scott, R. L., & Stone, D. A. (1986). MMPI profile constellations in incest families. Journal of Consulting and Clinical Psychology, 54, 364-368.

Scott, R. L., & Tetreault, L. A. (1987). Attitudes of rapists and other violent offenders toward women. Journal of Social Psychology, 127, 375-380.

Segal, Z. V., & Marshall, W. L. (1985). Self-report and behavioral assertion in two groups of sexual offenders. Journal of Behavior Therapy and Experimental Psychiatry, 16, 223-229.

Segal, Z. V., & Stermac, L. (1984). A measure of rapists' attitudes towards women. International Journal of Law and Psychiatry, 7, 437-440.

Simkins, L., Ward, W., Bowman, S., & Rinck, C. M. (1989). The Multiphasic Sex Inventory: Diagnosis and prediction of treatment response in child sexual abusers. Annals of Sex Research, 2, 205-226.

Smith, R. L., & Bradley, D. W. (1980). In defense of the Attitudes Toward Women Scale: An affirmation of validity and reliability. Psychological Reports, 47, 511-522.

Smith, W. R., & Monastersky, C. (1986). Assessing juvenile sexual offenders' risk for reoffending. Criminal Justice and Behavior, 13, 115-140.

Smith, M. D., & Self, G. D. (1981). Feminists and traditionalists: An additudinal comparison. Sex Roles, 7, 183-188.

Spence, J. T., & Helmreich, R. L. (1978). Masculinity and femininity: Their psychological dimensions, correlates and antecedents. Austin: University of Texas Press.

Spence, J. T., & Helmreich, R. L. (1972). The Attitudes Towards Women Scale: An objective instrument to measure attitudes towards the rights and roles of women in contemporary society. Psychological Documents, 2, (Ms. No. 153).

Spence, J. T., Helmreich, R. L., & Stapp, J. (1973). A short version of the Attitudes Toward Women Scale (AWS). Bulletin of the Psychonomic Society, 2, 219-220.

Stermac, L. E., Glancy, G., & Davidson, A. (1989, November). Relapse prevention for sexual aggression: Profiling risks for offenders. Presentation at the Annual Meeting of the American Association for the Advancement of Behavior Therapy, Washington, D.C.

Stermac, L. E., & Segal, Z. V. (1989). Adult sexual contact with children: An examination of cognitive factors. Behavior Therapy, 20, 573-584.

Sutherland, E. H. (1947). Principles of criminology (4th ed.). Chicago: J. B. Lippincott.

Taylor, L. (1972). The significance and interpretation of replies to motivational questions: The case of sex offenders. Sociology, 6, 23-39.

Templeman, T. L., & Stinnett, R. D. (1991). Patterns of sexual arousal and history in a "normal" sample of young men. Archives of Sexual Behavior, 20, 137-150.

Thorne, F. C. (1966a). A factorial study of sexuality in adult males. Journal of Clinical Psychology, 22, 375-386.

Thorne, F. C. (1966b). The sex inventory. Journal of Clinical Psychology, 22, 367-374.

Thorne, F. C., & Haupt, T. D. (1966). The objective measurement of sex attitudes and behavior in adult males. Journal of Clinical Psychology, 22, 395-403.

Toobert, S., Bartelme, K. F., & Jones, E. S. (1959). Some factors related to pedophilia. International Journal of Social Psychiatry, 4, 272-279.

Troldahl, V. C., & Powell, F. A. (1965). A short-form dogmatism scale for use in field studies. Social Forces, 44, 211-214.

Velasquez, R. J., Callahan, W. J., & Carrillo, R. (1989). MMPI profiles of Hispanic-American inpatient and outpatient sex offenders. Psychological Reports, 65, 1055-1058.

Whalley, L. J., & McGuire, R. J. (1978). Measuring sexual attitudes. Acta Psychiatrica Scandinavica, 58, 299-314.

Wilson, G. D. (1978). The secrets of sexual fantasy. London: J. M. Dent & Sons.

Wormith, J. S. (1984). Attitude and behavior change of correctional clientele: A three year follow-up. Criminology, 22, 595-618.

Wrightman, L. (1964). Measurement of philosophies of human nature. Psychological Reports, 14, 743-751.

Zuckerman, M. (1973). Scales for sex experience for males and females. Journal of Consulting and Clinical Psychology, 41, 27-29.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. Psychological Bulletin, 99, 434-442.